

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-311661

(43)Date of publication of application : 28.11.1995

(51)Int.Cl.

G06F 3/08

G06F 3/06

G06F 12/16

(21)Application number : 06-103927

(71)Applicant : FUJITSU LTD

(22)Date of filing : 18.05.1994

(72)Inventor : KURIHARA TAKUYA
SUGASAWA YASUYOSHI
MURAYAMA TAKASHI
NISHI HIDETOSHI

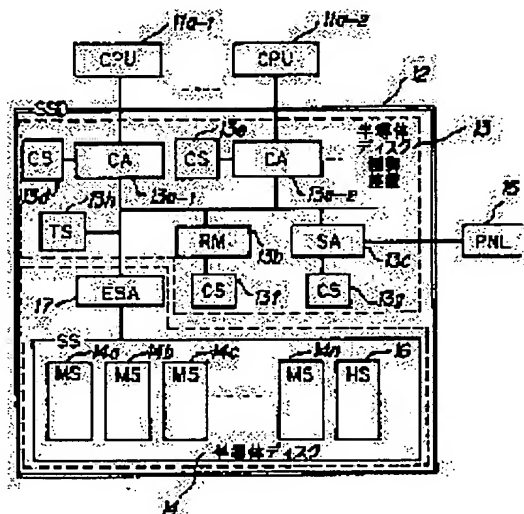
(54) SEMICONDUCTOR DISK DEVICE

(57)Abstract:

PURPOSE: To replace a semiconductor memory module without requiring a large-scale device and making a stop.

CONSTITUTION: Respective semiconductor memory modules 14a-14n are divided into plural access control units and a service adapter 13c copies the storage contents of the semiconductor module 14a where a memory error occurs to a stand-by semiconductor memory module 16 by the access control units.

Consequently, the storage contents of the semiconductor memory module where the memory error occurs can be saved without a stop only by adding the stand-by semiconductor memory module, so a resource manager 13b can perform exclusive control in the control units and host devices 11a-1-11a-2 can access the access control unit parts which are not the object of copying even during the copying process, thereby exerting no adverse influence on the access by the host devices.



LEGAL STATUS

[Date of request for examination]

26.12.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision]

of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-311661

(43) 公開日 平成7年(1995)11月28日

(51) Int.Cl.⁴
 G 0 6 F 3/08
 3/06
 12/16

識別記号
 H
 3 0 4 Z
 3 1 0 M

庁内整理番号
 7808-5B

F I

技術表示箇所

審査請求 未請求 請求項の数24 O L (全 36 頁)

(21) 出願番号 特願平6-103927

(22) 出願日 平成6年(1994)5月18日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中1015番地

(72) 発明者 栗原 拓弥

神奈川県横浜市港北区新横浜2丁目4番19号 株式会社富士通プログラム技研内

(72) 発明者 菅沢 康良

神奈川県横浜市港北区新横浜2丁目4番19号 株式会社富士通プログラム技研内

(72) 発明者 村山 孝

神奈川県横浜市港北区新横浜2丁目4番19号 株式会社富士通プログラム技研内

(74) 代理人 弁理士 斉藤 千幹

最終頁に続く

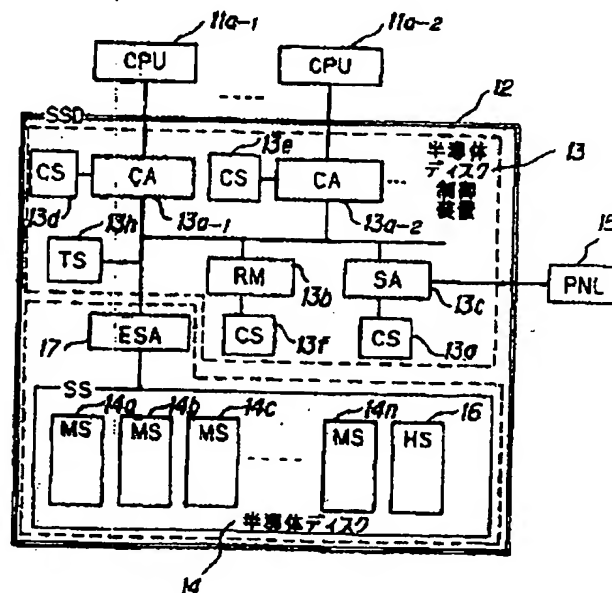
(54) 【発明の名称】 半導体ディスク装置

(57) 【要約】

【目的】 大がかりな装置を必要とせず、無停止で半導体メモリモジュールの交換を行う。

【構成】 各半導体メモリモジュール14a~14nをそれぞれ複数のアクセス制御単位に分け、サービスアダプタ13cは該アクセス制御単位でメモリエラーが生じた半導体メモリモジュール14aの記憶内容を予備の半導体メモリモジュール16にコピーする。これにより、予備の半導体メモリモジュール16を追加するだけで、メモリエラーが生じた半導体メモリモジュールの記憶内容を無停止で退避することができる。また、アクセス制御単位でコピーを行うためリソースマネージャ13bは該制御単位で排他制御を行うことができ、コピー中であっても上位装置11a-1~11a-2はコピー対象でないアクセス制御単位部分にアクセスすることができ、上位装置のアクセスに悪影響を与えることがない。

本発明の半導体ディスク装置の構成図
 (第1実施例)



【特許請求の範囲】

【請求項1】 複数の半導体メモリモジュールと、
予備の半導体メモリモジュールと、
半導体メモリモジュールへのデータの書き込み及び読み出しを制御するチャンネルアダプタと、
半導体メモリモジュールへのアクセスに関して排他制御を実行する排他制御部と、
各半導体メモリモジュールのメモリエラーを監視し、メモリエラーが発生した半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーする制御を行うサービスアダプタを備え、
各半導体メモリモジュールをそれぞれ複数のアクセス制御単位に分け、サービスアダプタは該アクセス制御単位でメモリエラーが生じた半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーすることを特徴とする半導体ディスク装置。

【請求項2】 前記排他制御部はアクセス制御単位で排他制御を行うことを特徴とする請求項1記載の半導体ディスク制御装置。

【請求項3】 アクセス制御単位毎にコピーが正常に完了したか否かを示すコピー管理テーブルを備え、チャンネルアダプタはデータリード時に該コピー管理テーブルを参照し、アクセス対象である領域が正常の場合にはアクセスし、異常の場合にはアクセスせずエラーとし、又、データライト時、アクセス対象の領域にデータを書き込むと共に、該領域に応じたコピー管理データを正常とすることを特徴とする請求項1又請求項2記載の半導体ディスク制御装置。

【請求項4】 少なくともチャンネルアダプタは、各半導体メモリモジュールのアクセス制御単位毎に論理アドレスと物理アドレスの対応を保持する制御テーブルを備え、
サービスアダプタはアクセス制御単位のコピー完了後に、制御テーブルにおける該アクセス制御単位の物理アドレスをコピー先の物理アドレスに変更するように制御することを特徴とする請求項1又は請求項2記載の半導体ディスク装置。

【請求項5】 少なくともチャンネルアダプタは、各半導体メモリモジュールのアクセス制御単位毎に論理アドレスと物理アドレスの対応を保持する制御テーブルを備え、
メモリエラーが発生して半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーするに先だって、サービスアダプタは、エラーが生じた半導体メモリモジュールの各論理アドレスにオリジナルな物理アドレスとコピー先の物理アドレスとが対応するように制御テーブルを変更すると共に、チャンネルアダプタに該半導体メモリモジュールへのデータ書き込みをオリジナルな物理アドレスとコピー先の物理アドレスが示す2つの記憶エリアに書き込むように指示し、エラーが生じた半

(2)

特開平7-311661

2

導体メモリモジュールの予備の半導体メモリモジュールへのコピー終了後に、エラーが生じた半導体メモリモジュールの各論理アドレスに前記コピー先の物理アドレスのみが対応するように制御テーブルを変更することを特徴とする請求項1又は請求項2記載の半導体ディスク装置。

【請求項6】 チャンネルアダプタは上位装置よりライト命令を受信した時、排他制御部にアクセス許可を要求し、

10 排他制御部はライト命令の対象である領域がコピー中の場合には、サービスアダプタにライト命令の発生を通知し、

サービスアダプタは該通知によりコピーを停止し、チャンネルアダプタは該領域にデータを書き込むことを特徴とする請求項5記載の半導体ディスク装置。

【請求項7】 サービスアダプタは、ライト命令のアクセス単位がコピーのアクセス単位よりも大きい場合には、当該ライト命令に書き込まれたアクセス単位領域のコピーを行わず、小さい場合にはコピーのアクセス単位で再度コピーすることを特徴とする請求項6記載の半導体ディスク装置。

【請求項8】 サービスアダプタは、ライト命令のアクセス単位がコピーのアクセス単位よりも大きい場合には、当該ライト命令に書き込まれたアクセス単位領域のコピーを行わず、小さい場合にはライト命令で書き込まれた領域以外の部分をコピーすることを特徴とする請求項6記載の半導体ディスク装置。

【請求項9】 複数の論理ドライブが割り付けられた半導体ディスクと、

30 半導体ディスクへのデータの書き込み及び読み出しを制御するチャンネルアダプタと、

半導体ディスクをバックアップすると共に、複数の論理ドライブが割り付けられたディスク装置と、

半導体ディスクに記憶されている各論理ドライブのデータをディスク装置の対応する論理ドライブに退避させると共に、ディスク装置の各論理ドライブに退避されているデータを半導体ディスクの対応する論理ドライブに復元するディスクアダプタとを備えた半導体ディスク装置において、

40 半導体ディスクにおける各論理ドライブの先頭アドレス及び容量を保持する第1の構成情報テーブルと、

ディスク装置における各論理ドライブの先頭アドレスと容量を保持する第2の構成情報テーブルと、

半導体ディスクの構成変更及びディスク装置に退避されているデータの半導体ディスクへの復元を指示する指示手段と、

構成変更指示に基づいて前記第1の構成情報テーブルが変更された場合、該第1の構成情報テーブルと変更前の第2の構成情報テーブルを用いて、論理ドライブ毎にディスクアドレスを半導体ディスクのアドレスに変換する

3

ためのアドレス変換テーブルを作成する手段を備え、ディスクアダプタは復元に際して該アドレス変換テーブルを用いてディスク装置の各論理ドライブに退避されているデータを半導体ディスクの対応する論理ドライブに復元する半導体ディスク装置。

【請求項10】 前記アドレス変換テーブルを作成する手段はディスクアダプタであり、ディスクアダプタは復元が指示された時にアドレス変換テーブルを作成する請求項9記載の半導体ディスク装置。

【請求項11】 前記第1の構成情報テーブルは半導体ディスクに記憶し、前記第2の構成情報テーブルはディスク装置に記憶する請求項10記載の半導体ディスク装置。

【請求項12】 ディスクアダプタは、復元完了後に前記第2の構成情報変換テーブルを第1の構成情報変換テーブルに整合させる請求項11記載の半導体ディスク装置。

【請求項13】 半導体メモリと、半導体メモリへのレコードの書き込み及び読み出しを制御する制御部を備え、磁気ディスク装置のトラックに相当するメモリ領域毎に、該メモリ領域に記録されているレコードの制御情報を記録するディレクトリを設け、該ディレクトリの内容に基づいて該メモリ領域にレコードを記録し、あるいはレコードを再生する半導体ディスク装置において、制御部はレコードに前記ディレクトリに記録される制御情報の一部を復元するために必要な制御情報復元データを含ませてメモリ領域に記録することを特徴とする半導体ディスク装置。

【請求項14】 前記制御情報は各レコードのトラック先頭からのオフセットアドレス、各レコードの所属するセクタのセクタ値を含む請求項13記載の半導体ディスク装置。

【請求項15】 前記制御情報の一部を復元するために必要なデータは、現レコードの先頭から次のレコードの先頭までのオフセットアドレス及び現レコードが所属するセクタのセクタ値である請求項14記載の半導体ディスク装置。

【請求項16】 上位装置より指示されたレコードの読み取りあるいは書き込み時にディレクトリにエラーが発生した場合、制御部は上位装置の指示により各レコード中に記録されている制御情報復元データを用いて上位装置より指示された前記レコードのサーチを行うことを特徴とする請求項13記載の半導体ディスク装置。

【請求項17】 上位装置より指示されたレコードの読み取りあるいは書き込み時にディレクトリにエラーが発生した場合、制御部は上位装置の指示により各レコード中に記録されている制御情報復元データを用いて上位装置より指示された前記レコードのサーチを行うと共に、各レコードの制御情報復元データを用いてディレクトリを再構築することを特徴とする請求項13記載の半導体

(3)

特開平7-311661

4

ディスク装置。

【請求項18】 上位装置より指示されたレコードの読み取りあるいは書き込み時にディレクトリにエラーが発生した場合、制御部は自動的に該ディレクトリに応じた各レコードに記録されている制御情報復元データを用いて上位装置より指示された前記レコードのサーチを行うことを特徴とする請求項13記載の半導体ディスク装置。

【請求項19】 上位装置より指示されたレコードの読み取りあるいは書き込み時にディレクトリにエラーが発生した場合、制御部は自動的に該ディレクトリに応じた各レコードに記録されている制御情報復元データに基づいて上位装置より指示された前記レコードのサーチを行うと共に、各レコードの制御情報復元データを用いてディレクトリを再構築することを特徴とする請求項13記載の半導体ディスク装置。

【請求項20】 少なくとも1トラック分のレコード及び該トラックのディレクトリを記憶するための予備の半導体メモリを備え、

コピー時、半導体メモリの所定のトラックのディレクトリにエラーが発生した場合、制御部は該トラックのレコードを予備の半導体メモリにコピーしながら該レコード中の制御情報復元データを用いてディレクトリを再構築し、再構築したディレクトリを予備の半導体メモリのディレクトリ領域に書き込むことを特徴とする請求項13記載の半導体ディスク装置。

【請求項21】 半導体ディスクと、上位装置からの指示に従って半導体ディスクへのレコードの書き込み及び読み出しを制御する制御部を備えた半導体ディスク装置において、

半導体ディスクは、レコードに含まれる制御情報を記憶する制御情報記憶領域と、多数のブロックに分割されると共に、各ブロックはレコードに含まれるデータを記憶するデータ記憶領域を備え、

制御部はレコードに含まれるデータを圧縮、復元する圧縮・復元機構を備え、

上位装置からの書き込みコマンドを受信した時、制御部はデータを圧縮してデータ記憶領域の所定ブロックに格納し、かつ、制御情報に該データが格納されるブロックのブロックアドレスを含ませて制御情報記憶領域に格納することを特徴とする半導体ディスク装置。

【請求項22】 各ブロックの使用／未使用を管理する半導体メモリ管理部を備え、

レコード書替え時、制御部は書替え前のレコードの制御情報を前記制御情報記憶領域より読み取って該レコードのデータが格納されているブロックの数 B_0 を求め、書替えデータを圧縮して得られる圧縮データの格納に必要なブロック数 B_n を求め、 $B_n=B_0$ の場合には圧縮データを前のデータが格納されていたブロックに書き込む請求項22記載の半導体ディスク装置。

5

【請求項23】 $B_n > B_0$ の場合には、制御部は不足分のブロックを前記半導体メモリ管理部に要求し、圧縮データを前のデータが格納されていたブロックと半導体メモリ管理部から指示されたブロックに書き込み、かつ、制御情報に該指示されたブロックのブロックアドレスを追加することを特徴とする請求項22記載の半導体ディスク装置。

【請求項24】 $B_n < B_0$ の場合には、制御部は余分のブロックを前記半導体メモリ管理部に通知し、圧縮データを前のデータが格納されていたブロックのうち半導体メモリ管理部に通知したブロック以外のブロックに書き込み、かつ、制御情報より半導体メモリ管理部に通知したブロックのブロックアドレスを削除することを特徴とする請求項22記載の半導体ディスク装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は半導体ディスク装置に係り、特に、磁気ディスク装置(DASD: Direct Access Storage Device)に格納されるユーザデータを全て半導体メモリに書き込むことにより機械動作を伴わずに上位装置から高速にアクセスできるようにした磁気ディスク装置に関する。

【0002】

【従来の技術】半導体ディスク装置は、磁気ディスク装置の振舞(コマンドコード、データ転送の仕方等)を維持したまま、記録媒体を磁気ディスクから半導体メモリに置き換えたものである。このため、上位装置(CPU)と半導体ディスク制御装置間のインターフェースはCPUと磁気ディスク制御装置とのインターフェースと全く同一になっている。この半導体ディスク装置によれば、磁気ディスク装置のようにヘッドの移動が不要のため、瞬時にアクセスできる利点があり、しかも、CPUと磁気ディスク制御装置間のソフトウェア資産をそのまま使える利点がある。

【0003】図57はかかる半導体ディスク装置の構成図である。1aはCPU、2は半導体ディスク装置(SSD: Shared Storage Device)、3は半導体ディスク制御装置、4は半導体ディスクであり、複数の半導体メモリモジュール(MS: Main Storage) 4a、4b、4c、...、4n及び半導体メモリモジュールへのデータの書き込み/読み出しを制御するメモリインターフェースアダプタ(ESP: Extended Storage Adaptor) 4sを備えている。5は保守パネルあるいはパソコンである。半導体ディスク制御装置3において、3aは上位装置(CPU) 1aとの間に単一あるいは複数のインターフェース(上位インターフェース)を有するチャネルアダプタCAであり、図では1つしか示してないが複数のチャネルアダプタが設けられている。3bは排他制御テーブル(図示せず)を備え、いずれの上位インターフェースも所定の半導体メモリモジュールを使用していない場合に

(4)

特開平7-311661

6

は他の上位インターフェースに該半導体メモリモジュールの使用を許可し、使用中の場合には使用を許可しない排他制御を実行するリソースマネージャRMである。尚、実際には、半導体メモリモジュールは複数の論理ドライブに分割されており、リソースマネージャRMは各論理ドライブ毎に排他制御を行う。3cは各ユニットのIML(イニシャル・マイクログラム・ローディング)処理や状態監視処理、障害時のリカバリ処理を行うサービスアダプタ(SA: Service Adaptor)である。3d、3e、3fは各種制御テーブルやプログラムを記憶する制御記憶部(Control Storage)である。

【0004】

【発明が解決しようとする課題】

・第1の課題

半導体ディスク装置においては、半導体メモリモジュールの障害は致命的である。半導体メモリモジュールに障害が発生すると、従来は、データを退避させ、しかる後、電源を切断し、障害が発生した半導体メモリモジュールを別の半導体メモリモジュールで交換する。交換後、電源を投入して半導体ディスク装置を立ち上げ、データを復元させるという手順をとっている。しかし、かかる方法では、データを退避させる装置が別途必要になり、しかも、電源切断時並びにデータ退避/復元時に半導体ディスク装置を使用できなくなり、ノンストップ装置という要求に答えられない問題がある。又、半導体ディスク装置を停止することなく障害が発生した半導体メモリモジュールを交換する方法も提案されている(特開平3-268020号公報、名称:半導体ディスクの無停止保守方式)。しかし、この提案されている半導体ディスクの無停止保守方式では、半導体ディスク装置とは別に、無停止保守を可能にするための大がかりな保守装置が必要となる問題がある。

【0005】・第2の課題

ところで、半導体ディスク装置は電源を切断するとその記憶内容が消失する。このため、半導体ディスク装置にバックアップディスク装置が接続される場合がある。半導体ディスクを構成する各半導体メモリモジュールは複数の論理ドライブに分割され、上位装置はStart I/O命令により論理ドライブを指定し、該論理ドライブが使用可能であれば該論理ドライブの所定位置にアクセスするようになっている。かかる半導体ディスクの論理ドライブの構成とバックアップディスク装置における論理ドライブの構成は1:1の関係がある。しかし、半導体ディスクの論理ドライブのサイズを変更したり、位置を変更すると、半導体ディスクとバックアップディスク装置における論理ドライブの構成が1:1に対応しなくなる。このため、半導体ディスクの構成を変更する前にバックアップディスク装置にデータを退避しておいても、半導体ディスク装置の論理ドライブの構成を変更すると、バックアップディスク装置に退避しておいたデータを半導

7

体ディスクに復元できなくなり、変更前のユーザデータが無効化されて使用できなくなる問題がある。

【0006】・第3の課題

半導体ディスク装置は、磁気ディスク装置をエミュレーションしているため、実デバイスが有するユーザデータ部に加えてディレクトリ(DIRECTORY)とよばれるコントロール情報部を各トラック毎に持ち、該ディレクトリにより現在エミュレートしているトラックフィールド内のレコード(ユーザデータ)に関するメモリ上のアドレスや各セクタ情報等を管理している。チャンネルアダプタは、指定されたトラックフィールドのユーザデータエリアをアクセスする場合、指定されたトラックのディレクトリをチャンネルアダプタ内部に取り込むことによって、目的のトラックフィールドの制御情報を取得し、その情報に従ってユーザデータをアクセスする。このディレクトリは、①当該トラックフィールドに書き込まれているレコードのうち最終レコードのレコード番号や②セクタディレクトリ、③レコードディレクトリから成っている。セクタディレクトリは、セットセクタ処理で何番目のレコードが最初に読み出せるかが書き込まれたテーブルであり、レコードディレクトリは、各レコード毎のトラック先頭からの相対アドレス(オフセットアドレス)が書き込まれたテーブルであり、直接レコード番号で位置付けたレコードフィールドをアクセスするために用いるものである。これら情報を用いることにより、セットセクタ処理やサーチID処理時に直接オリエント(仮想ヘッド位置)を移動することができ、アクセスを高速に行うことができる。

【0007】ところで、メモリの2ビットエラー等で目的のディレクトリ部を読み込めなくなると、当該ディレクトリが制御するトラックフィールドのユーザデータが全てアクセス不可能になってしまい、事実上、有効なユーザデータが失われてしまう。このため、データ消失回避機構として、ディレクトリが読み込み不可能な状態に陥った場合でも、ユーザデータをアクセスできるようにするデータ消失回避機構が必要になる。従来の半導体ディスク装置では、チャンネルアダプタのファームウェアが指定するトラックフィールドをアクセスしようとした時、まず当該トラックフィールドの制御情報が書き込まれているディレクトリを読み込んでから、通常処理を開始する。そのため、ディレクトリが読み込み不可能な場合、当該トラックフィールドへのアクセスをしようとしても、ディレクトリ崩壊時の回復手段がないため、アクセスが不可能である。そこで、読み込み不可能となったトラックを再度利用できるようにするために、最低単位のイニシャライズ(例えば1シリンダ単位)をして、崩壊したディレクトリを初期化し、ディレクトリを有効化してやる必要がある。しかし、イニシャライズ処理はイニシャライズする領域のユーザデータを消失させてしまうので、あらかじめイニシャライズする領域のデータを

(5)

特開平7-311661

8

退避しておかなければならない。又、ディレクトリの崩壊によってアクセス不可能に陥ったトラックフィールドのデータはチャンネルコマンドの1つとして用意されたメモリダンプ命令でしか読みだすことができない。このため、半導体ディスク装置内部のトラックフォーマットを知っている者以外は、当該トラックフィールドのユーザデータとして必要な部分を抜き出して再現することは不可能であり、完全に元のデータを復旧させることは殆ど不可能であった。

10 【0008】・第4の課題

半導体ディスク装置は、データの記憶・格納媒体が半導体メモリチップである。このため、磁気ディスク装置に比較すると、ビット当りの記憶コストが高くなる。又、半導体ディスク装置1台当りの記憶容量は少なくなる。この容量に関する問題を解決するために、データを圧縮して書き込み、読み出す場合には圧縮データを復元する方法が提案されている。かかる圧縮方式の問題点は、ユーザデータを書き替える際に、圧縮後のサイズが圧縮前のサイズと異なる場合が生じることである。圧縮後のデータサイズが小さい場合には、メモリの有効利用のために余分な領域を解放する必要があり、逆に圧縮後のデータサイズが大きい場合には新たな領域を確保して書き替える必要がある。かかるメモリ管理は煩雑であり、従来は、メモリの有効利用を図りつつ、しかも簡単な方法でメモリ領域の解放、割当て管理を行うことができない問題があった。

【0009】以上から本発明の第1目的は、別途保守装置を設けなくても半導体ディスク装置を運用中のまま、半導体メモリモジュールの交換ができ、しかも、半導体メモリモジュール交換中におけるアクセス速度を低下しないようにできる半導体ディスク装置を提供することである。本発明の第2の目的は、半導体ディスクの構成を変更しても、構成変更前にバックアップディスク装置に退避したデータを構成変更後の半導体ディスク装置に復元でき、変更前のユーザデータが使用することができる半導体ディスク装置を提供することである。本発明の第3の目的は、ディレクトリを読み込めなくなっても、ディレクトリを用いずに指定された目的レコードへのアクセスが可能であり、しかも、該ディレクトリの再構築が可能な半導体ディスク装置を提供することである。本発明の第4の目的は、データを圧縮して半導体メモリに書き込む場合、効果的にメモリ領域の解放、割当てを行ってメモリの有効利用を図ることができる半導体ディスク装置を提供することである。

【0010】

【課題を解決するための手段】上記第1の課題は、本発明によれば、複数の半導体メモリモジュールと、予備の半導体メモリモジュールと、上位装置からのアクセスコマンドに従って半導体メモリモジュールへのデータの書き込み及び読み出しを制御する制御部(チャンネルアダプ

(6)

特開平7-311661

9

タ)と、各半導体メモリモジュールのメモリエラーを監視し、メモリエラーが発生した半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーする制御を行うサービスアダプタとにより達成される。

【0011】上記第2の課題は、本発明によれば、複数の論理ドライブが割り付けられた半導体ディスクと、半導体ディスクをバックアップすると共に、複数の論理ドライブが割り付けられたディスク装置と、半導体ディスクに記憶されている各論理ドライブのデータをディスク装置の対応する論理ドライブに退避させると共に、ディスク装置の各論理ドライブに退避されているデータを半導体ディスクの対応する論理ドライブに復元するディスクアダプタと、半導体ディスクにおける各論理ドライブの先頭アドレス及び容量を示す第1の構成情報テーブルを保持する手段と、ディスク装置における各論理ドライブの先頭アドレスと容量を示す第2の構成情報テーブルを保持する手段と、半導体ディスクの構成変更及びディスク装置に退避されているデータの半導体ディスクへの復元を指示する指示手段と、構成変更指示に基づいて前記第1の構成情報テーブルが変更された場合、該第1の構成情報テーブルと変更前の第2の構成情報テーブルを用いて、論理ドライブ毎にディスクアドレスを半導体ディスクのアドレスに変換するためのアドレス変換テーブルを作成する手段とにより達成される。

【0012】上記第3の課題は、本発明によれば、半導体メモリと、半導体メモリへのレコードの書き込み及び読み出しを制御すると共に、レコードにディレクトリに記録される制御情報の一部を復元するために必要な制御情報復元データを含ませて半導体メモリに記録する制御部(チャンネルアダプタ)により達成される。上記第4の課題は、本発明によれば、ディレクトリやレコードに含まれる制御情報を記憶する制御情報記憶領域と多数のブロックに分割されると共に、レコードに含まれるデータ部を記憶するデータ記憶領域を備えた半導体ディスクと、上位装置からの指示に従って半導体ディスクへのレコードの書き込み及び読み出しを制御すると共にレコードに含まれるデータ部を圧縮、復元する圧縮・復元機構を備えた制御部(チャンネルアダプタ)と、各ブロックの使用/未使用を管理する半導体メモリ管理部とにより達成される。

【0013】

【作用】第1の発明においては、各半導体メモリモジュールをそれぞれ複数のアクセス制御単位に分ける。サービスアダプタは該アクセス制御単位でメモリエラーが生じた半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーする。このようにすれば、予備の半導体メモリモジュールを追加するだけで、メモリエラーが生じた半導体メモリモジュールの記憶内容を無停止で退避することができ、大がかりな保守装置を必要としない。また、アクセス制御単位でコピーを行うため

10

該制御単位で排他制御を行うことができ、コピー中でも上位装置はコピー対象でないアクセス制御単位部分にアクセスすることができ、上位装置のアクセスに悪影響を与えることがない。

【0014】第2の発明においては、半導体ディスクの構成変更後に、指示手段よりディスク装置に退避されているデータの半導体ディスクへの復元を指示する。この復元が指示された時、アドレス変換テーブル作成手段は、半導体ディスクの構成情報テーブルとディスク装置の構成情報テーブルとを用いて、論理ドライブ毎にディスクアドレスを半導体ディスクのアドレスに変換するためのアドレス変換テーブルを作成する。ディスクアダプタは復元に際して、該アドレス変換テーブルを用いてディスク装置の各論理ドライブに退避されているデータを半導体ディスクの対応する論理ドライブに復元する。以上のようにすれば、半導体ディスクの構成を変更しても、構成変更前にバックアップ用のディスク装置に退避したデータを構成変更後の半導体ディスクに正しく復元でき、変更前のユーザデータが使用することができ、

【0015】第3の発明において、チャンネルアダプタは、ディレクトリの内容に基づいて半導体メモリへのレコードの書き込み及び読み出しを制御する。この場合、チャンネルアダプタは前記ディレクトリに記録される制御情報の一部を復元するために必要な制御情報復元データをレコードに含ませて記録する。このようにすれば、チャンネルアダプタは、ディレクトリを読み込めなくなっても、該ディレクトリを用いずに各レコードに含まれる制御情報復元データを用いて上位装置により指示されたレコードへのアクセスが可能となり、しかも、該ディレクトリの再構築ができる。

【0016】第4の発明においては、半導体ディスクを、①ディレクトリやレコードに含まれる制御情報を記憶する制御情報記憶領域と②レコードのデータ部を記憶するデータ記憶領域に分け、該データ記憶領域を多数のブロックに分割する。上位装置からの書き込みコマンドを受信した時、チャンネルアダプタは、データ部を圧縮してデータ記憶領域の所定ブロックに格納し、かつ、制御情報に該データが格納されるブロックのブロックアドレスを含ませて圧縮せずに制御情報記憶領域に格納する。このように、ユーザデータを圧縮して記憶するからメモリの有効利用が図れ、しかも、制御情報は圧縮しないため、復元処理が不要となり上位装置のデータアクセス時間の短縮を図ることができる。

【0017】又、第4の発明において、レコード書替え時、チャンネルアダプタは書替え前のレコードのデータ部が格納されているブロックの数 B_0 を求めると共に、書替えデータを圧縮して得られる圧縮データの格納に必要なブロック数 B_1 を求め、 $B_1 = B_0$ の場合には圧縮データを前のデータ部が格納されていたブロックに書き込む。一方、 $B_1 > B_0$ の場合には、チャンネルアダプタは不

34

(7)

特開平7-311661

11

12

足分のブロックを半導体メモリ管理部に要求し、圧縮データを前のデータ部が格納されていたブロックと半導体メモリ管理部から指示されたブロックに書き込む。更に、 $B_n < B_0$ の場合には、チャネルアダプタは余分のブロックを半導体メモリ管理部に通知し、圧縮データを前のユーザデータが格納されていたブロックのうち半導体メモリ管理部に通知したブロック以外のブロックに書き込む。このようにすれば、データ部を圧縮して半導体メモリに書き込む場合、効果的にメモリ領域の解放、割当てを行ってメモリの有効利用を図ることができる。

【0018】

【実施例】

(a) 本発明の第1の実施例

(a-1) 半導体ディスク装置の構成

図1は本発明の半導体ディスク装置の構成図である。11a-1, 11a-2, ...はCPU(上位装置)、12は半導体ディスク装置(SSD)、13は半導体ディスク制御装置、14は半導体ディスク、15はサービスアダプタ(後述)に各種指示を出して保守を実行させるメンテナンスパネル(PNL)である。半導体ディスク14は、複数のユーザデータ格納用の半導体メモリモジュール(MS: Main Storage) 14a, 14b, 14c, ... 14nと、半導体メモリモジュールにエラーが発生した時、該半導体メモリモジュールが新たな半導体メモリモジュールと交換されるまでその代役を勤める予備の半導体メモリモジュール(HS: Hot Spare Memory) 16、各半導体メモリモジュールへのデータの書き込み/読み出しを制御するメモリインターフェースアダプタ(ESP: Extended Storage Adaptor) 17を備えている。予備の半導体メモリモジュール(HS)は1枚しか示していないが、2枚以上設けることもできる。

【0019】各半導体メモリモジュールは所定サイズ、例えばCPUからのアクセスの排他制御単位であるシリンダ単位で多数の範囲(範囲1~範囲n)に区分されており、メモリエラー発生時にはこの範囲毎に予備の半導体メモリモジュール16への退避制御が行われ、該範囲のサイズがアクセス制御単位となる。例えば、半導体メモリモジュール14aにエラーが発生した時の交換手順は、大略以下ようになる。

①エラーが発生した半導体メモリモジュール14aの記憶内容をアクセス制御単位に予備の半導体メモリモジュール(HS) 16に退避する。排他制御はこのアクセス制御単位で行われる。

②退避完了後、予備の半導体メモリモジュール16がユーザデータ格納用の半導体メモリモジュールとなる。

③全アクセス制御単位の退避完了後に、メモリエラーが発生した半導体メモリモジュール14aを新たな半導体メモリモジュールと交換する。以後、この交換された半導体メモリモジュール14aが予備の半導体メモリモジュールとなる。尚、半導体メモリモジュール16を予備

に戻すためには、①の手順と同様に半導体メモリモジュール16の記憶内容をアクセス制御単位に半導体メモリモジュール14aに復元する。

【0020】半導体ディスク制御装置13において、13a-1, 13a-2, ...は上位装置(CPU) 11a-1, 11a-2, ...との間に単一あるいは複数のインターフェース(上位インターフェース)を有するチャネルアダプタ(CA)、13bはリソースマネージャ(RM)であり、排他制御テーブルECT(図示せず)を備え、いずれのユニットも所定の半導体メモリモジュールのある領域を使用していない場合には、要求により他のユニットに半導体メモリモジュールの該領域へのアクセスを許可し、使用中の場合にはアクセスを許可しない排他制御を実行する。尚、本実施例では、リソースマネージャはアクセス制御単位で使用/未使用を管理し、該アクセス制御単位で排他制御を行う。

【0021】13cは各ユニットのIML(イニシャル・マイクロプログラム・ローディング)処理や状態監視処理、障害時のリカバリ処理、例えばメモリエラー発生時における半導体メモリモジュールの交換処理を行うサービスアダプタ(SA)である。13d~13gは各種制御テーブルCTLやプログラムを記憶する制御記憶部(CS)である。制御テーブルCTLには、図2に示すように各半導体メモリモジュールの範囲(アクセス制御単位)毎に論理アドレス(CCHH)と半導体ディスクの物理アドレスの対応が記憶されている。論理アドレスとはCPUより指定されるアドレス(例えば磁気ディスクのデータアドレスで指定される場合は、シリンダ/ヘッド番号CCHH)、物理アドレスとは半導体ディスク装置内部で半導体メモリモジュールをアクセスする実アドレスである。各モジュールは制御テーブルCTLを用いて論理アドレスを物理アドレスに変換する。

【0022】13hはコピーの成功/不成功(有効/無効)を示すコピー管理テーブルCCTを記憶する記憶部(TS)である。このコピー管理テーブルCCTは半導体メモリモジュールに書き込まれるデータのうちの壊れた単位でコピーの成功/不成功を管理する。コピーが失敗した場合(媒体エラーで読めなかった場合等)、当該部分を無効とする。例えば、図3に示すようにトラック毎にコピーの成功/不成功を管理する。磁気ディスクをエミュレーションレートしている半導体ディスクの場合、シリンダ(アクセス制御単位)毎にコピーしていた場合に、エラーしてもそのシリンダ全てが無効ではない。そこで、コピー管理テーブルCCTによりトラック毎にエラーしたトラックのみ無効にすれば、きめ細かいデータの保証ができる。この場合、トラック毎でなくレコード単位で管理するとより細かいデータの保証ができる。コピー管理テーブルCCTは各モジュールが参照/更新できる場所におけばよく、半導体メモリモジュール内に配置することもできる。

13

【0023】(a-2) 本発明の半導体メモリの第1の交換制御の概略

半導体ディスク14は、図4に示すように3枚の半導体メモリモジュール(MS①~MS③)14a~14cと、1枚の予備の半導体メモリモジュール(HS)16で構成され、それぞれ5個のアクセス制御単位で構成されている。初期時、各半導体メモリモジュール14a~14cの制御テーブルCTLの内容は図5に示すようになっている。尚、図中、CTL aは半導体メモリモジュール14a(MS①)の制御テーブル部分、CTL bは半導体メモリモジュール14b(MS②)の制御テーブル部分、CTL cは半導体メモリモジュール14c(MS③)の制御テーブル部分である。又、図中、XXXXは論理アドレスを、MS①-00は半導体メモリモジュール14a(MS①)のオフセットアドレス00(図4参照)を意味しており、他の表記も同様である。

【0024】半導体メモリモジュール14b(MS②)にメモリエラーが発生すると、サービスアダプタSAの制御でアクセス制御単位で半導体メモリモジュール14b(MS②)から予備の半導体メモリモジュール16にコピーが行われる。図6に示すように、コピー処理が進行して、第1、第2の制御単位⑥、⑦のコピーが終了し、第3の制御単位⑧のコピーは他のモジュールにより使用中のため行えず、現在第4の制御単位⑨をコピー中で、最後の制御単位をまだコピーしてないとすると、その時点で制御テーブルCTLの内容は図7に示すようになる。すなわち、コピーが完了した制御単位(範囲)⑥⑦の物理アドレスはコピー先の半導体メモリモジュール16の物理アドレスとなる。以後、制御単位毎のコピーが行われ、全ての制御単位のコピーが完了すれば、制御テーブルCTLの内容は図8に示すようになる。すなわち、コピーが完了した全ての制御単位(範囲)の物理アドレスはコピー先の物理アドレスとなる。

【0025】(a-3) 本発明の半導体メモリモジュールの交換制御処理
・コピー処理

図9及び図10は半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーするためのコピー処理のフロー図である。保守パネル(PNL)15からの指示により、あるいはエラー検出管理により、サービスアダプタ(SA)13cはメモリエラー発生時の半導体メモリモジュール(半導体メモリモジュール14bとする)を認識する(ステップ101)。尚、サービスアダプタSAは常時各メモリモジュールから1アドレスずつ順次データを読み出してメモリエラーが発生していないか監視している(パトロール)。1ビットエラーは元のデータに復元できるため、直ちにメモリエラー発生としないが設定値以上の1ビットエラーを検出した時メモリエラー発生と認定する。又、2ビットエラーはメモリーインターフェースアダプタ(ESP)17よりサービス

(8)

特開平7-311661

14

アダプタSAに通知される。

【0026】サービスアダプタSAはメモリエラー発生時の半導体メモリモジュールを認識すると、リソースマネージャ(RM)13bに予備の半導体メモリモジュール(HS)16の使用許可を要求する(ステップ102)。これは予備の半導体メモリモジュールHSが他の半導体メモリモジュールMSの交換作業に使用されていないことを確認するためのものである。例えば、半導体ディスク制御装置13が二重化されている場合において、他方のサービスアダプタSAが交換作業を行っている場合があるからである。リソースマネージャRMは半導体メモリモジュールHSが使用中の場合にはサービスアダプタSAに不許可通知を出す(ステップ103、104)。これにより、サービスアダプタSAは、使用許可になる迄待つ。一方、半導体メモリモジュールHSが使用中でなければ、リソースマネージャRMは使用許可をサービスアダプタSAに通知し、内蔵の排他制御テーブルECTに「半導体メモリモジュールHSが使用中になったこと」を記入する(ステップ105)。ついで、サービスアダプタSAはリソースマネージャRMにエラー発生時の半導体メモリモジュールの第1番目のアクセス制御単位部分の使用許可を要求する(ステップ106)。

【0027】使用可能であれば、リソースマネージャRMはサービスアダプタSAに使用許可を通知すると共に、排他制御テーブルECTに「第1番目のアクセス制御単位が使用中になったこと」を記入する(ステップ107、108)。サービスアダプタSAは使用許可通知により、当該制御単位のコピーを実行する(ステップ109)。コピー中にエラーが発生せず、当該アクセス制御単位のコピーが完了すれば(ステップ110、111)、サービスアダプタSAは各モジュール13a-1、13a2、13b等に制御テーブルCTLの変更を指示する(ステップ112)。すなわち、サービスアダプタSAは、制御テーブルCTL bの制御単位⑨における物理アドレスをHS-00に変更するように各モジュールに指示する。各モジュールは指示された通りに制御メモリCSの制御テーブルCTLを変更し、変更完了をサービスアダプタSAに通知する(ステップ113)。尚、サービスアダプタSAも制御メモリCSの制御テーブルCTLを変更する。

【0028】サービスアダプタSAは全モジュールから変更完了通知を受信すれば、リソースマネージャRMにアクセス終了を通知する。これによりリソースマネージャRMは、「第1番目のアクセス制御単位が不使用になったこと」を記入する(ステップ114)。ついで、サービスアダプタSAは半導体メモリモジュール14bの全制御単位のコピーが終了したチェックする(ステップ115)。全制御単位のコピーが終了すれば、サービスアダプタSAはチャネルアダプタCAを介して上位装置

15

CPUにコピー終了を通知すると共に、保守パネルPNLにコピー完了を表示し（ステップ116）、コピー処理を終える。コピー処理が完了すれば、保守員がメモリエラーを生じた半導体メモリモジュール（MS）14bを新たな半導体メモリモジュールと交換する。一方、全制御単位のコピーが終了していなければ、コピー対象をコピーを終了していない別の制御単位に変更し（ステップ117）、以後ステップ106以降の処理を繰り返す。ステップ107において、第1番目のアクセス制御単位が使用中である場合には、リソースマネージャRMはサービスアダプタSAに使用不許可通知を出す（ステップ118）。これにより、サービスアダプタSAはステップ115に飛び、以降の処理を繰り返す。

【0029】又、ステップ110において、コピー時に所定トラックにおいてエラーが発生すれば、サービスアダプタSAはチャンネルアダプタCAを介して上位装置CPUにエラー発生を通知すると共に（ステップ119）、コピー管理テーブルCCTの該当トラックにコピー不成功（無効）を記入する（ステップ120）。以後、ステップ111以降の処理を繰り返す。図11はコピーの成功／不成功を示すコピー管理テーブルCCTの説明図であり、制御単位⑥（半導体メモリモジュール14bの第1アクセス制御単位）のコピー管理テーブル部分を示している。尚、制御単位⑥は3つのトラック1～3で構成されているものとしている。図11(a)は制御単位⑥の全トラックのコピーがエラーを生じることなく成功した場合であり、コピー管理テーブルCCTのトラック1～3には「有効」が記入される。図11(b)はトラック2のコピー時にエラーが発生した場合であり、コピー管理テーブルCCTのトラック2に「無効」が記入され、トラック1、3には「有効」が記入される。尚、コピーが不成功のトラックにデータを書き込んだ場合には、該トラックは「有効」になり、コピー管理テーブルCCTが書き替えられる。

【0030】・半導体メモリモジュール交換後の処理
図12は半導体メモリモジュール交換後の処理フロー図である。コピー完了後に保守員がエラーを生じた半導体メモリモジュールを交換すれば、サービスアダプタSAは交換完了を感知し（ステップ131）、以後、交換後の半導体メモリモジュールを予備の半導体メモリモジュールとする（ステップ132）。尚、交換完了後、保守員の保守パネルPNLからの指示により、半導体メモリモジュール16を予備の半導体メモリモジュールに戻すように構成することもできる。又、交換完了後、自動的に半導体メモリモジュール16を予備の半導体メモリモジュールに戻すように構成することもできる。図13はかかる場合の処理フロー図である。サービスアダプタSAは半導体メモリモジュールの交換完了を感知し（ステップ141）、前述のコピー処理と同様の手順で半導体メモリモジュール16の記憶データを交換後の半導体メ

(9)

特開平7-311661

16

モリモジュール14bにコピーする（ステップ142）。

【0031】・コピー管理テーブルを用いたアクセス処理

各モジュールは半導体メモリモジュールよりレコードを読み出す場合にはコピー管理テーブルCCTを参照する。図14はコピー管理テーブルCCTを参照したアクセスの処理フロー図である。まず、レコード読み出し、書き込みかを判断し（ステップ151）、読み出しの場合には、アクセス先のトラックの有効／無効をコピー管理テーブルCCTを参照して調べ（ステップ152）、有効の場合には該レコードを読み出し（ステップ153）、無効の場合にはエラー処理を行う（ステップ154）。一方、書き込みの場合には、アクセス先のトラックが無効であれば「有効」に変更し（ステップ155）、しかる後、該トラックにレコードを書き込む（ステップ156）。

【0032】以上のように、予備の半導体メモリモジュールを追加するだけで、メモリエラーが生じた半導体メモリモジュールの記憶内容を無停止で退避することができ、大がかりな保守装置を必要としない。また、半導体メモリモジュールをアクセス制御単位に細分化し、該アクセス制御単位でコピーを行うため、該制御単位で排他制御を行うことができ、コピー中であっても上位装置はコピー対象でない他のアクセス制御単位部分にアクセスすることができ、上位装置のアクセスに悪影響を与えることがない。又、アクセス制御単位のコピーが完了する毎にサービスアダプタSAは、チャンネルアダプタCA等に設けられている制御テーブルCTLにおける物理アドレスをコピー先の物理アドレスに変更するため、アクセス制御単位のコピー完了後、該アクセス制御単位については直ちにコピー先（予備の半導体メモリモジュール）にアクセスすることができる。更に、アクセス制御単位毎にコピーが正常に完了したか否かを管理するコピー管理テーブルを設け、チャンネルアダプタCA等のモジュールはデータリード時に該コピー管理テーブルを参照し、アクセス対象である領域（トラック）が正常の場合にはアクセスし、異常の場合にはアクセスせずエラーとする。この結果、コピー時にコピーエラーが発生しても誤ったデータを読み取って処理することがなく、誤動作を防止できる。

【0033】(a-4) 本発明の半導体メモリの第2の文換制御の概略

以上は制御テーブルCTLが図2に示す構成を備えている場合である。制御テーブルCTLは図15に示すように構成することもできる。この制御テーブルCTLは、各半導体メモリモジュールの範囲（アクセス制御単位）毎に、①論理アドレス（CCHH）と、②半導体ディスクのオリジナルな物理アドレスと、③コピー先の物理アドレスを記入する欄を備えている。論理アドレスはCD

(10)

特開平7-311661

17

U等より指定されるアドレス（例えば磁気ディスクのデータアドレスで指定される場合は、シリンドラ／ヘッド番号CCHH）、オリジナルな物理アドレスは半導体ディスク装置内部で半導体メモリモジュールをアクセスする実アドレスである。通常は、論理アドレスに対応させてオリジナルな物理アドレスのみが対応付けされており、コピー先の物理アドレスは無効（マークを付けるなどして制御）となっている。

【0034】半導体ディスク14は、図4に示すように3枚の半導体メモリモジュール（MS①～MS③）14a～14cと、1枚の予備の半導体メモリモジュール（HS）16で構成され、それぞれ5個のアクセス制御単位で構成されているものとする。初期時、各半導体メモリモジュール14a～14cの制御テーブルCTLの内容は図16に示すようになっている。尚、図中、CTL aは半導体メモリモジュール14a（MS①）の制御テーブル部分、CTL bは半導体メモリモジュール14b（MS②）の制御テーブル部分、CTL cは半導体メモリモジュール14c（MS③）の制御テーブル部分である。又、図中、XXXXは論理アドレスを、MS①-00は半導体メモリモジュール14a（MS①）のオフセットアドレス00を意味しており、他の表記も同様である。

【0035】半導体メモリモジュール14b（MS②）にメモリエラーが発生すると、サービスアダプタSAの制御でアクセス制御単位で半導体メモリモジュール14b（MS②）から予備の半導体メモリモジュール16にコピーが行われる。この場合、コピーに先だって、サービスアダプタSAは各モジュールに対し制御テーブルの変更を指示する。すなわち、エラーを生じた半導体メモリモジュール14bの各論理アドレスに、①元の物理アドレスと②コピー先の物理アドレスとがそれぞれ対応するように制御テーブルの変更を指示する。これにより、各モジュールの制御テーブルCTLは図17に示すように変更される。又、サービスアダプタSAは、チャンネルアダプタCA等のモジュールに対して、半導体メモリモジュール14bへのアクセスに際してフォーキング処理を行うように指示する。フォーキング処理とは、図18に示すように、データ読み出しはオリジナルな物理アドレスが指示するエリアからデータを読み出し、データ書き込みは、オリジナルな物理アドレスとコピー先の物理アドレスが示す2つの記憶エリアに書き込む処理である。

【0036】以後、図19に示すようにアクセス制御単位のコピーを行う。サービスアダプタSAは半導体メモリモジュール14bの全アクセス制御単位のコピーが終了すれば、各モジュールに制御テーブルCTLのオリジナルな物理アドレスとコピー先の物理アドレスの交換を指示する。これにより、各モジュールの制御テーブルCTLは図20に示すように変更される。以後、サービスアダプタSAは、各モジュールに対してフォーキング処

18

理の終了を通知すると共に、コピー先の物理アドレス記入欄のアドレスを無効にするように指示してコピー処理を終了する。この結果、各モジュールの制御テーブルCTLは図21に示すようになり、オリジナルな物理アドレスはコピー先の物理アドレスとなる。

【0037】(a-5) 本発明の半導体メモリモジュールの交換制御

・コピー処理

図22及び図23は半導体メモリモジュールの記憶内容を予備の半導体メモリモジュールにコピーするためのコピー処理のフロー図である。尚、制御テーブルCTLは図16に示すように構成されているものとする。保守パネル（PNL）15からの指示により、あるいはエラー管理により、サービスアダプタ（SA）13cはメモリエラー発生時の半導体メモリモジュール（半導体メモリモジュール14bとする）を認識する（ステップ201）。サービスアダプタSAはメモリエラー発生時の半導体メモリモジュールを認識すると、リソースマネージャ（RM）13bに予備の半導体メモリモジュール（HS）16の使用許可を要求する（ステップ202）。リソースマネージャRMは半導体メモリモジュールHSが使用中の場合にはサービスアダプタSAに不許可通知を出す（ステップ203、204）。これにより、サービスアダプタSAは、使用許可になる迄待つ。一方、半導体メモリモジュールHSが使用中でなければ、リソースマネージャRMは使用許可をサービスアダプタSAに通知し、内蔵の排他制御テーブルECTに「半導体メモリモジュールHSが使用中になったこと」を記入する（ステップ205）。

【0038】ついで、サービスアダプタSAはチャンネルアダプタCA等の各モジュールに半導体メモリモジュール14bの交換開始及びフォーキング処理開始を指示する（ステップ206）。又、サービスアダプタSAは各モジュールに対し制御テーブルの変更を指示する。これにより、各モジュールは指示された通りに制御テーブルCTLを変更し（図17参照）、変更終了をサービスアダプタSAに通知する（ステップ207）。各モジュールから制御テーブルの変更終了が通知されると、サービスアダプタSAはリソースマネージャRMにエラー発生時の半導体メモリモジュールの第1番目のアクセス制御単位の使用許可を要求する（ステップ208）。

【0039】使用可能であれば、リソースマネージャRMはサービスアダプタSAに使用許可を通知すると共に、排他制御テーブルECTに「第1番目のアクセス制御単位が使用中になったこと」を記入する（ステップ209、210）。サービスアダプタSAは使用許可通知により、当該制御単位のコピーを実行する（ステップ211）。コピー中にエラーが発生せず、当該アクセス制御単位のコピーが完了すれば（ステップ212、213）、サービスアダプタSAはリソースマネージャRM

17

(11)

特開平7-311661

19

20

にアクセス終了を通知する。これによりリソースマネージャRMは、「第1番目のアクセス制御単位が不利用になったこと」を記入する(ステップ214)。ついで、サービスアダプタSAは半導体メモリモジュール14bの全制御単位のコピーが終了したチェックする(ステップ215)。全制御単位のコピーが終了していなければ、コピー対象をコピーを終了していない別の制御単位に変更し(ステップ216)、以後ステップ208以降の処理を繰り返す。ステップ209において、第1番目のアクセス制御単位が使用中である場合には、リソースマネージャRMはサービスアダプタSAに使用不許可通知を出す(ステップ217)。これにより、サービスアダプタSAはステップ215に飛び、以降の処理を繰り返す。

【0040】ステップ212において、コピー時に所定トラックにおいてエラーが発生すれば、サービスアダプタSAはチャネルアダプタCAを介して上位装置CPUにエラー発生を通知すると共に(ステップ218)、コピー管理テーブルCCTの該当トラックにコピー不成功(無効)を記入する(ステップ219)。以後、ステップ213以降の処理を繰り返す。ステップ215において、半導体メモリモジュール14bの全制御単位のコピーが終了した場合には、サービスアダプタSAは各モジュール13a-1、13a-2、13b等に対して、半導体メモリモジュール14bに関係する制御テーブル部分のオリジナルな物理アドレスとコピー先の物理アドレスの交換を指示する(ステップ220)。これにより、各モジュールは制御テーブルCTLを指示通りに変更し(図20参照)、変更終了をサービスアダプタSAに通知する(ステップ221)。ついで、サービスアダプタSAは、各モジュールに対してコピー完了及びフォーキング処理の終了を通知すると共に(ステップ222)、半導体メモリモジュール14bに関係する制御テーブル部分のコピー先物理アドレス欄を無効にするように指示する(ステップ223)。各モジュールは制御テーブルCTLを指示された通りに変更する(図21参照)。これにより、オリジナルな物理アドレスはコピー先の物理アドレスとなる。

【0041】以後、サービスアダプタSAはチャネルアダプタCAを介して上位装置CPUにコピー終了を通知すると共に、保守パネルPNLにコピー完了を表示し(ステップ224)、コピー処理を終える。コピー処理が完了すれば、保守員がメモリエラーを生じた半導体メモリモジュール(MS)14bを新たな半導体メモリモジュールと交換する。以上のように、予備の半導体メモリモジュールを追加するだけで、メモリエラーが生じた半導体メモリモジュールの記憶内容を無停止でコピーすることができる。又、チャネルアダプタ等のモジュールはコピー中の半導体メモリモジュールにデータを書き込む際、フォーキング処理に従ってオリジナルな物理アド

レスとコピー先の物理アドレスが示す2つの記憶エリアにデータを書き込むようにしたから、第1の交換制御処理のように制御単位のコピー完了毎に制御テーブルを変更する必要がなく、コピーを高速に行うことができる。

【0042】コピー中の制御単位エリアにライト命令が発生した場合の処理

コピー中の制御単位エリアに上位装置CPUからライト命令が発生する場合がある。かかる場合には、ライト命令をコピーが完了するまで待たせることもできるが上位装置のアクセス速度が低下する。そこで、ライト命令を優先してコピーを一時停止させるように制御する。図24はかかるコピー中の制御単位エリアにライト命令が発生した場合の処理の流れ図である。上位装置CPUよりライト命令がチャネルアダプタCAに発行されると(ステップ251)、チャネルアダプタはアクセス制御単位についてリソースマネージャRMにアクセス許可を要求する(ステップ252)。リソースマネージャは、該アクセス制御単位がコピー中であるか調べ、コピー中でなければチャネルアダプタCAにアクセス許可を通知する(ステップ253、254)。これにより、チャネルアダプタCAはライトコマンドをフォーキング処理に従って実行する。

【0043】一方、コピー中の場合には、リソースマネージャ255はサービスアダプタSAにコピー中の制御単位エリアに上位装置よりライトコマンドが発生したことを通知する(ステップ255)。これにより、サービスアダプタSAはコピーを停止する(ステップ256)。ついで、リソースマネージャRMはチャネルアダプタCAにアクセスを許可し(ステップ257)、チャネルアダプタCAはライトコマンドをフォーキング処理に従って実行する。チャネルアダプタCAによる書き込みが完了すれば(ステップ258)、サービスアダプタSAは前記書き込みのアクセス単位(トラック数=A_w)とコピーのアクセス単位(トラック数=A_cで例えばA_c=1)の大小を比較する(ステップ259)。A_w≧A_cの場合には、サービスアダプタSAは書き込みコマンドのアクセス単位についてコピーが完了したものとみなし(ステップ260)、以後残りの部分についてコピーを再開する。

【0044】一方、A_w<A_cの場合には、書き込みコマンドのアクセス単位を包含するコピー単位は、コピーが完了していないものとみなして再度コピーし、あるいは、アクセス単位以外のコピー単位部分をコピーし(ステップ261)、以後残りの部分についてコピーを再開する。以上のように、コピーを停止して上位装置CPUの書き込みコマンドを実行するようにしたから上位装置のアクセス速度は低下しない。又、アクセス単位がコピー単位より大きい場合には、アクセス単位部分のコピーは完了したものとみなせるからコピー速度を向上することができる。

21

【0045】(a-6) 半導体ディスク装置の実際の構成
図25は半導体ディスク装置の全体構成図であり、二重化構成になっており、添字1を有するモジュールは第1の半導体ディスク装置G0側のモジュールであり、添字2を有するモジュールは第2半導体ディスク装置G1側のモジュール、添字の無いモジュールは共通のモジュールである。尚、制御記憶部CSは各モジュールに内蔵されている。CAは上位装置のチャンネルとのインターフェース制御を行うチャンネルアダプタであり、電気チャンネル、光チャンネル、OCリンクに対応する種々のチャンネルアダプタが適宜接続されるようになっている。RMはリソースマネージャであり排他制御、論理バス管理等の処理動作を制御すると共にサブシステム全体の資源管理を行う。SAはサービスアダプタであり、自分がマスターになり、他のユニット(モジュール)の状況を管理する。

【0046】C-BUSは各ユニットがメッセージ通信及び制御情報のアクセスを行う制御用バス、D-BUSは各ユニットが半導体ディスクとの間でデータを授受するデータ転送用バス、S-BUSはサービスモジュールがマスターになり、各ユニットの状況を管理するサービスバスである。BH-1、BH-2はバスの競合制御、バスクロックの分配を行うバスハンドラ、MDKはメモリ障害時に該メモリの内容を一時的にバックアップする磁気ディスク装置(オプション)、DAは磁気ディスク装置とのインターフェース制御を行うデバイスアダプタ、BANKは半導体ディスク(共用メモリ)であり、半導体メモリモジュールMS、HS及び予備の半導体メモリモジュールHSが搭載されるようになっている。ESP1~ESP4は半導体ディスクへのアクセス制御を行うポート(Extended Storage Port)、ESA1~ESA4はESPと半導体メモリモジュールMS間のタイミング制御や、メモリのリフレッシュ、エラーチェックコードに基づくデータ修正を実行するメモリインタフェースアダプタ、PA-NELは保守パネルである。

【0047】第1、第2半導体ディスク装置G0、G1は中央点線を中心に対称に構成されており、上位CPUは対称に第1、第2半導体ディスク装置のチャンネルアダプタCA1、CA2に接続され、それぞれのポートESP2、ESP3は他方のメモリアダプタESA3、4と接続されている。従って、一方のチャンネルアダプタに障害が生じてCPUは他方のチャンネルアダプタから半導体ディスクをアクセスすることができる。又、一方の半導体ディスクが障害を生じて他方の半導体ディスクにアクセスすることができ、信頼性を向上している。チャンネルアダプタCA、リソースマネージャRM、サービスアダプタSA等のモジュールはそれぞれマイクロプロセッサで構成されており、おおむね図26に示す構成を備えている。図において、91はマイクロプロセッサ(MPU)、92はRAM構成の制御記憶部(CS)、93は

(12)

特開平7-311661

22

ROM構成の制御記憶部(CS)、94は内部バスに接続されたドライバ/レシーバ(DV/RV)、95はバスインターフェースロジック(BIL)、96は外部インターフェースと接続されたドライバ/レシーバ(DV/RV)、97はバッファ又はテーブル記憶部(TS)、98は個別LSI(ゲートアレイ)である。尚、接続される外部インターフェースの数によってドライバ/レシーバ(DV/RV)96の数が異なる。

【0048】(b) 本発明の第2の実施例

10 (b-1) 全体の構成

半導体ディスク装置は電源を切断するとその記憶内容が消失する。このため、半導体ディスク装置にバックアップディスク装置が接続される場合がある。図27はバックアップディスク装置を備えた半導体ディスク装置の構成図であり、図1と同一部分には同一符号を付している。11a-1はCPU(上位装置)、12は半導体ディスク装置(SSD)、13は半導体ディスク制御装置、14は半導体ディスク、15はサービスアダプタSAに各種指示を出して保守を実行させるメンテナンスパネル(PNL)である。

【0049】半導体ディスク14は、複数のユーザデータ格納用の半導体メモリモジュール(MS: Main Storage)14a、14b、14c・・・と、半導体メモリモジュールにエラーが発生した時、該半導体メモリモジュールが新たな半導体メモリモジュールと交換されるまでその代役を勤める予備の半導体メモリモジュール(11S)16を備えている。尚、17はメモリインタフェースアダプタである。半導体ディスク制御装置13において、13a-1はチャンネルアダプタ(CA)、13bはリソースマネージャ(RM)、13cはサービスアダプタ(SA)、13d、13f、13gは各種制御テーブルCTLやプログラムを記憶する制御記憶部(CS)である。18は電源切断時等において半導体ディスク14に記憶されているデータを記憶するバックアップ用の磁気ディスク装置、19は磁気ディスク装置18に半導体ディスク14に記憶されているデータを退避すると共に、退避されたデータを磁気ディスク装置より読み出して半導体ディスク14に復元するディスクアダプタである。

【0050】(b-2) 半導体ディスクと磁気ディスクの論理ドライブ構成

半導体ディスク14を構成する各半導体メモリモジュール(MS)14a~14cはそれぞれ1枚のプリント板で構成され、メモリバンクのスロットに挿入されている。各半導体メモリモジュール(MS)14a~14cはそれぞれ所定サイズの複数の論理ドライブに分割されており、上位装置CPUはStar: I/O命令により所定の論理ドライブを指定し、該論理ドライブが使用可能であれば該論理ドライブの所定位置にアクセスするようになっている。図28は半導体ディスク14と磁気ディスク装置18の論理ドライブの構成図である。半導体ディス

23

ク14は2枚の半導体メモリモジュール14a、14bで構成されているものとしており、第1半導体メモリモジュール14aは3つの論理ドライブ0～3を備え、それぞれのサイズは10シリンダとなっている。又、第2の半導体メモリモジュール14bは10シリンダサイズの2つの論理ドライブ3～4を備えている。

【0051】磁気ディスク装置18における論理ドライブの構成は矢印で示すように半導体ディスク14の論理ドライブ構成と1:1の対応関係がある。すなわち、磁気ディスク装置18の論理ドライブ構成と半導体ディスク14の論理ドライブ構成は同一になっている。半導体ディスク14における論理ドライブの構成を示す第1の構成情報テーブル21は、論理ドライブ0～4毎に先頭アドレス（シリンダアドレス）及び容量（シリンダ数）を保持しており、所定の半導体メモリモジュール（マスターモジュール）14aに格納されている。又、ディスク装置18における論理ドライブの構成を示す第2の構成情報テーブル22は、第1の構成情報テーブル21と同一の構成を有し磁気ディスク装置18に記憶されている。

【0052】半導体ディスク14に記憶されているデータを磁気ディスク装置18に退避させる場合には、ディスクアダプタ（DA）19は第1、第2の構成情報テーブル21、22を読み取る。ついで、該第1の構成情報テーブル21を参照して第i論理ドライブ（ $i=0\sim4$ ）のデータを半導体ディスク14より読み取り、該データを第2の構成情報テーブル22を参照して磁気ディスク装置18の第i論理ドライブ（ $i=0\sim4$ ）に格納する。又、磁気ディスク装置18に退避されているデータを半導体ディスク14に復元させる場合には、ディスクアダプタ（DA）19は第1、第2の構成情報テーブル21、22を読み取る。ついで、該第2の構成情報テーブル22を参照して第i論理ドライブ（ $i=0\sim4$ ）のデータを磁気ディスク18より読み取り、該データを第1の構成情報テーブル21を参照して半導体ディスク14の第i論理ドライブ（ $i=0\sim4$ ）に格納する。

【0053】(b-3) 半導体ディスク14の論理ドライブ構成を変更した場合の問題点

ところで、半導体ディスク14のデータを磁気ディスク装置18に退避した後に、半導体ディスク14の論理ドライブ構成を変更したい場合がある。図29は論理ドライブ構成の変更説明図であり、(a)は変更前の構成、(b)は変更後の構成であり、論理ドライブ0のサイズを10シリンダから20シリンダに増加し、論理ドライブ1～論理ドライブ4の先頭アドレスを変更した場合である。半導体ディスク14の構成を前記のように変更すると、従来装置では半導体ディスク14と磁気ディスク18の第1、第2の構成情報テーブル21、22は図30に示すようになる。この結果、磁気ディスク18に退避されているデータを半導体ディスク14に復元すると、矢印

(13)

特開平7-311661

24

に示すように復元され、正しくデータの復元ができなくなる。すなわち、磁気ディスク18に退避されている論理ドライブ0、1のデータ（点線で囲んである）が半導体ディスク14の論理ドライブ0に復元され、論理ドライブ2のデータが論理ドライブ1に復元され、論理ドライブ3のデータが論理ドライブ2に復元され、論理ドライブ4のデータが論理ドライブ3に復元されてしまう。この結果、退避したデータを使用できなくなる。以上は、論理ドライブのサイズを変更した場合であるが、図31に示すように論理ドライブの配置を変更した場合も同様の問題点が生じる。

【0054】(b-4) 本発明の復元処理の概略

図32はデータ退避後に半導体ディスク14の論理ドライブ構成を変更した場合における本発明の復元処理の概略説明図である。14は半導体ディスク、18は磁気ディスク装置19はディスクアダプタである。半導体ディスク14において、20はユーザデータ格納領域、21は半導体ディスクの論理ドライブ構成を保持する第1構成情報テーブル、23は後述するアドレス変換テーブルである。磁気ディスク装置18において、22は磁気ディスクの論理ドライブ構成を保持する第2構成情報テーブル、24はユーザデータ格納領域である。半導体ディスク14及び磁気ディスク18は初期時に図28に示す構成を備えている。半導体ディスク14のデータを磁気ディスク18に退避した後、保守パネル（PNL）15からの指示により半導体ディスク14の構成を図29(b)に示すように変更する。この結果、半導体ディスク14の第1の構成情報テーブル21は図30に示すようになる。

【0055】ついで、保守パネルPNLより復元が指示されると、ディスクアダプタ（DA）19は、半導体ディスク14及び磁気ディスク18より第1、第2の構成情報テーブル21、22を読み取り、これら第1、第2の構成情報テーブル21、22を用いて、論理ドライブ毎にディスクアドレスを半導体メモリのアドレスに変換するためのアドレス変換テーブル23（図33参照）を作成する。このアドレス変換テーブル23は論理ドライブ毎に①半導体ディスク14の先頭アドレス、シリンダ数、及び②磁気ディスク18の先頭アドレス、シリンダ数を持たせたものである。しかる後、ディスクアダプタDAは図34に示すようにアドレス変換テーブル23を参照して第i論理ドライブ（ $i=0\sim4$ ）のデータを磁気ディスク18より読み取り、該データをアドレス変換テーブル23が指示する半導体ディスク14の第i論理ドライブ（ $i=0\sim4$ ）に格納する。以上により、磁気ディスク18に退避してある各論理ドライブのデータを構成変更後の論理ドライブに復元でき、該データを使用することができる。復元完了後に、磁気ディスク装置18の第2の構成情報テーブル22を半導体ディスク14の第1の構成情報テーブル21に一致させる。

(14)

特開平7-311661

25

【0056】(b-5) 本発明の復元処理

図35は本発明の復元処理のフロー図である。データを磁気ディスク装置18に退避した後に、半導体ディスク14の論理ドライブ構成を変更したい場合には、保守パネル(PNL)15より構成情報変更コマンドと論理ドライブの構成情報を入力する(ステップ301)。これにより、サービスアダプタSAは半導体ディスク14に記憶されている第1の構成情報テーブル21を更新する(ステップ302)。しかる後、磁気ディスク装置18に退避してあるデータを半導体ディスク14へ復元するために、保守パネルPNLよりリストアコマンドを入力する(ステップ303)。サービスアダプタSAはリストアコマンドが入力されると、ディスクアダプタ(DA)19に対して、アドレス変換テーブル23の作成を指示する(ステップ304)。

【0057】ディスクアダプタDAは半導体ディスク14及び磁気ディスク装置18より第1、第2の構成情報テーブル21、22を読み取り、これら第1、第2の構成情報テーブル21、22を用いてアドレス変換テーブル23(図32参照)を作成し、半導体ディスク14に記憶する(ステップ305)。ついで、ディスクアダプタDAは該アドレス変換テーブル23を参照して第i論理ドライブ(i=0~4)のデータを磁気ディスク18より読み取り、該データをアドレス変換テーブル23が指示する半導体ディスク14の第i論理ドライブ(i=0~4)に格納する(ステップ306、307)。

【0058】全データの半導体ディスク14への復元が完了すれば、保守パネルPNLより構成情報合わせ込みコマンドを発行する(ステップ308)。サービスアダプタSAは構成情報合わせ込みコマンドが発行されると、ディスクアダプタDAに第2の構成情報テーブル22の変更を指示する(ステップ309)。これにより、ディスクアダプタDAは半導体ディスク14より第1の構成情報テーブル21を読み出し、第1、第2の構成情報テーブル21、22が1:1に対応するように第2の構成情報テーブル22を作成し、磁気ディスク18に記憶されている内容を更新する(ステップ310)。以上より、半導体ディスクの構成を変更しても、構成変更前にバックアップディスク装置に退避したデータを構成変更後の半導体ディスク装置に正しく復元でき、変更前のユーザデータが使用することができる。

【0059】(c) 本発明の第3の実施例

(c-1) トラックフォーマット、ディレクトリ及びカウン

ト部の構成

・トラックフォーマット
半導体ディスク装置では、各トラックフィールドの先頭にディレクトリ(コントロール情報部)が設けられている。チャンネルアダプタCAは指定されたトラックフィールドのユーザデータをアクセスする場合、該トラックフィールドのディレクトリを読み込み、該ディレクトリを

26

用いてユーザデータをアクセスする。図36は半導体ディスクにおけるトラックフォーマットの例であり、各トラックの先頭にはディレクトリ(DIRECTORY)が配置され、その後にホームアドレス(HA)が置かれ、以降にCKDフォーマットにしたがった可変長のレコード(RECORD-0, RECORD-1, ...)が配列されている。ホームアドレスHAはトラックアドレス(CCHH)を示すものであり、各レコードはカウント部C、キー部K、データ部Dで構成されている。

10 【0060】・ディレクトリ

ディレクトリは図37に示すように448バイトを有し、①1バイトの最終レコード番号(LAST RECORD NO.)、②218バイトのセクタディレクトリ(SECTOR DIRECTORY)、③190バイトのレコードディレクトリ(RECORD DIRECTORY)、④その他有効性を確認するためのID部等のエリア、⑤未使用エリアで構成されている。最終レコード番号(LAST RECORD NO.)はトラックフィールドに書き込まれている最後のレコードのレコード番号を示すものである。各トラックフィールドの最終レコードの直後にはエンドマーク(EOF)が書き込まれており、最終レコードはこのエンドマークを読み込む直前のレコードということになる。そして、最終レコードのレコード番号が最終レコード番号と成る。

【0061】セクタディレクトリ(SECTOR DIRECTORY)は、セクタ処理で最初に読める物理レコード番号をセクタ値別に格納したものである。物理レコード番号とは、ホームアドレスHAを0として、論理レコード番号に1を加えた値である。トラックは例えば218のセクタに区分され、それぞれセクタ値0~217迄の値が割り当てられる。セクタ値217はトラックを所定サイズのセクタで分割した時の最後のセクタを示すもので、最大セクタ値である。このセクタディレクトリ(SECTOR DIRECTORY)はセクタ命令に対するセクタ処理において使用され、セクタ処理で何番目のレコードが最初に読み出せるかを示す。尚、実際は0~221のセクタが存在する。但し218~221はレコード配置の関係上これ以降にレコードがくることはないのでセクタディレクトリには存在しない。レコードディレクトリ(RECORD DIRECTORY)は各レコード(レコード0~レコード94)までのトラック先頭からのオフセットアドレスを2バイトで示すものである。トラックの先頭アドレスに目的レコードのオフセットアドレスを加えることにより、該目的レコードのカウント部のメモリアドレスを求めることができる。レコード番号94はトラックに書き込める最大のレコード数を示す。このレコードディレクトリ(RECORD DIRECTORY)は、セクタ処理で取得したレコード番号のフィールドの読み出しや直接レコード番号を指定してレコードを読み出す際に使用される。

40 【0062】・カウント部

各レコードは前述のように、カウント部C、検索のため

27

のキーが記録されるキー部K(必ずしも必要でない)、ユーザデータが記録されるデータ部Dで構成されている。カウント部Cは図38に示すように64バイトを有し、①ホストから与えられる8バイトの実データ部、②データの保証のために前後に付加される各4バイトのID部、③トラックフィールドをアクセスするための48バイトのコントロール情報部より成っている。実データ部には、トラックアドレス(CCHH)、レコード番号(R)、及び後に続くキー部の長さ(KL)、データ部の長さ(DL)が記録される。ID部にはレコードの先頭を識別するコードや、データが書き込まれた最新の時刻(タイムスタンプ)等が記録される。コントロール情報部は、ディレクトリが消失しても目的レコードを検索し、かつ、ディレクトリの一部を復元できるようにするためにディレクトリ復元データ(制御情報復元データ)が記録されている。ディレクトリ復元データには、①自分のレコードの先頭から次のレコード先頭迄の相対アドレス(オフセットアドレス)Aiと②自分のレコードが位置するセクタのセクタ値Siが含まれている。

【0063】ディレクトリ、ホームアドレス、各レコードのカウント部の長さは一定のため、ディレクトリ、HA及びRECORD-0のカウント部はトラック先頭から決まった相対位置にある。図39(a)に示すようにトラック先頭からホームアドレスHIAまでのオフセットアドレスLi0、レコードRi-1迄のオフアドレスをLiとすると、これらオフセットアドレスLi(i=0,1,2,...)はレコードディレクトリ(RECORD DIRECTORY)に記録される(図37参照)。この場合、レコードRi-1からレコードRi迄のオフセットアドレスをAiとすると、該オフセットアドレスAiはディレクトリ復元データとしてレコードRiのカウント部に記録される。Ai, Li間には次式
$$Li + Ai \rightarrow Li+1$$
 の関係が成立する。

【0064】又、セクタとレコードの関係が図39(b)に示すようになっているものとする、セクタディレクトリ(SECTOR DIRECTORY)におけるセクタ値0に対応させて0(HA)が、セクタ値1~2に対応させて1(レコードR0)が、セクタ値3~6に対応させてレコード2(R1)が、セクタ値7~13に対応させて3(レコードR2)が...記録される。すなわち、セクタiにレコードRjが存在すれば、セクタ値iに対応させてレコードRj+1がセクタディレクトリに記録される。そして、レコードRjがセクタi~(i+m)に位置すると、セクタ値i~(i+m)がディレクトリ復元データとしてレコードRj+1のカウント部に記録される。

【0065】(c-2)レコード読み取り、書き込み手順 図40は半導体ディスク装置のレコード読み取りシーケンス説明図である。尚、半導体ディスク装置は図1と同一構成を備えているものとする。上位装置CPUよりシークコマンドSKが発生すると、チャンネルアダプタCA

(15)

特開平7-311661

28

は直ちに動作終了信号を上位装置に返す。上位装置CPUは動作終了信号を受信すれば、つぎにセットセクタコマンドSSを発行する。以後、同様にして上位装置はサーチIDコマンドSIDを発行し、チャンネルアダプタCAはこれらコマンドを受信する。チャンネルアダプタCAはサーチIDコマンドを受信すると、シークコマンドで指示されている論理アドレス(トラックアドレスCCHH)を半導体ディスクの物理アドレスに変換する。この物理アドレスはトラックの先頭アドレス、換言すればディレクトリの先頭アドレスを示しているから、半導体ディスクよりディレクトリを読み込む。

【0066】ついで、セットセクタコマンドで指示されているセクタ値に応じたレコード番号をセクタディレクトリより求める(セットセクタ動作)。しかる後、該レコード番号のアドレスをレコードディレクトリより求め、該アドレスよりレコードのカウント部を読み取る。ついで、該カウント部に含まれるレコードのCCHHRとサーチIDコマンドで指示されている目的レコードのCCHHRが一致するかチェックし、一致すればサーチID動作を終了する。一致しない場合には次のレコード番号のレコードを読み取って同様の処理を行う。目的のレコードが見つければ、チャンネルアダプタCAは上位装置に動作終了信号を返す。これにより上位装置はリードコマンド(READ CKD)を発行し、チャンネルアダプタCAは該コマンドを受信する。以後、チャンネルアダプタCAはカウント部C、キー部K、データ部Dを順次上位装置CPUに転送し、最後に正常終了を送って一連のレコードリード処理を終了する。

【0067】図41は半導体ディスク装置のレコード書き込みシーケンス説明図であり、サーチ動作迄はレコード読み取り手順と同一である。サーチID動作により目的とするレコードが見つければ、チャンネルアダプタCAは上位装置に動作終了信号を返す。これにより上位装置はライトコマンド(WRITE CKD)を発行し、チャンネルアダプタCAは該コマンドを受信する。以後、チャンネルアダプタCAは、ディレクトリ復元データ(制御情報復元データ)Ai, Siを求めてカウント部に挿入し、しかる後、レコードを記録する。レコード記録完了後、チャンネルアダプタCAは上位装置CPUに書き込み終了を通知する。通知後、ディレクトリの内容を更新し、ディレクトリが記録されていた半導体ディスク位置に書き戻す。

【0068】(c-3)ディレクトリの読み込みエラーが発生した場合の処理

①スキップサーチ処理

レコードの読み込み、書き込み時に、チャンネルアダプタCAがディレクトリを読み込めなくなった場合には、各レコードのカウント部に記録されているディレクトリ復元データを用いて目的とするレコードをサーチ(スキップサーチ)する。尚、ディレクトリが読み込めなくなった場合としては、2ビットエラーの発生、IDエラーの

29

発生（ディレクトリの書き込み失敗）等があり、かかるエラー発生時、チャンネルアダプタCAはディレクトリを取り込むことができず、通常のアクセスができなくなる。図42はチャンネルアダプタCAがディレクトリを読み込めなくなった場合のスキップサーチ処理のフロー図、図43はスキップサーチ処理の説明図である。まず、チャンネルアダプタは $1 \rightarrow i$ とする（ステップ401）。ついで、チャンネルアダプタCAはアドレス $Li-1$ （ $=L0$ ）よりレコード $Ri-1$ （ $=R0$ ）のカウンタ部C0を読み込む（ステップ402）。なお、レコードR0のトラック先頭からの記憶位置は固定でL0は既知である。

【0069】カウンタ部の読み込み後、該カウンタ部に含まれるレコード番号がアクセス指示されたレコード番号Nと等しいかチェックする（ステップ403）。等しければ、レコード $Ri-1$ が目的とするレコードであり、該レコードを読み込み、あるいは書替えを行う（ステップ404）。しかし、カウンタ部に記録されているレコード番号がアクセス指示されたレコード番号Nと等しくなければ、カウンタ部に記録されている次のレコード Ri 迄の相対アドレス（オフセットアドレス） Ai を取得する（ステップ405）。ついで、次式

$$Li-1 + Ai \rightarrow Li$$

により、レコード Ri のトラック先頭からのオフセットアドレス Li を演算し（ステップ406）、 Li 演算後 i を歩進し（ $i+1 \rightarrow i$ 、ステップ407）、以後ステップ402以降の処理を繰り返す。以上により、最終的にステップ403においてカウンタ部に記録されているレコード番号がアクセス指示されたレコード番号Nと等しくなり、目的とするレコードが求まり、該レコードの読み込み、あるいは書替えを行う。以上のようにすれば、上位装置CPUより指示されたレコードの読み取りあるいは書き込み時に、ディレクトリを読み込めなくなってもチャンネルアダプタCAは自動的に各レコード中に記録されているディレクトリ復元データを用いて上位装置より指示された目的レコードをサーチして該レコードの読み取りあるいは書き込みを行うことができる。

【0070】②ディレクトリの再構築

ディレクトリを読み込めなくなった場合には、上記スキップサーチ法により目的とするレコードをサーチすることができる。この場合、スキップサーチ処理と並行してディレクトリ復元データを用いてディレクトリを再構築することができる。

・レコードディレクトリの再構成処理

図44はレコードディレクトリの再構成処理フロー図である。まず、ホームアドレスHA、レコードR0のオフセットアドレスとしてL0、L1（固定値）をレコードディレクトリにセットする（ステップ421）。しかる後、 i を1にセットし、レコード $Ri-1$ のカウンタ部を読み込む（ステップ422、423）。ついで、読み込

(10)

特開平7-311661

30

んだ情報がエンドマーク（EOF）でないかチェックし（ステップ424）、EOFでなければ、カウンタ部より次のレコード Ri 迄の相対アドレス（オフセットアドレス） Ai を取得する（ステップ425）。

【0071】 Ai が求まれば、次式

$$Li-1 + Ai \rightarrow Li$$

により、レコード Ri のトラック先頭からのオフセットアドレス Li を演算する（ステップ426）。ついで、 Li をレコード Ri の先頭からのオフセットアドレスとしてレコードディレクトリにセットし（ステップ427）、 i を歩進し（ $i+1 \rightarrow i$ 、ステップ428）、以後ステップ423以降の処理を繰り返す。以上の処理を繰り返せば、いつかステップ424においてエンドマークがEOFが検出される。エンドマークが検出されればレコードディレクトリにおける残りを“FF”で埋める（ステップ425）。以上によりレコードディレクトリを復元することができる。

【0072】・セクタディレクトリの再構成処理

図45はセクタディレクトリの再構成処理フロー図である。ホームアドレスHA、レコードR0のカウンタ部の位置は固定である。従って、セクタディレクトリにおける最初の幾つかのセクタ値に対応させてH0、R0をセットする（ステップ451）。ついで、 i を1にし、レコード $Ri-1$ のカウンタ部を読み込む（ステップ452、453）。しかる後、カウンタ部より次のレコード Ri 迄の相対アドレス（オフセットアドレス） Ai を取得する（ステップ454）。

【0073】 Ai が求まれば、次式

$$Li-1 + Ai \rightarrow Li$$

により、レコード Ri のトラック先頭からのオフセットアドレス Li を演算する（ステップ455）。ついで、レコード Ri のカウンタ部を読み込み（ステップ456）、読み込んだ情報がエンドマーク（EOF）でないかチェックする（ステップ457）。EOFでなければ、レコード Ri のカウンタ部より該レコード Ri の先頭セクタ値 Si を取得する（ステップ458）。ついで、セクタディレクトリにおけるセクタ値 Si に対応させてレコード Ri を記入し（ステップ459）、しかる後、 i を歩進し（ $i+1 \rightarrow i$ 、ステップ460）、ステップ454以降の処理を繰り返す。

【0074】以上の処理を繰り返せば、いつかステップ424においてエンドマークがEOFが検出される。エンドマークが検出されればセクタディレクトリにおける残りを“FF”で埋める（ステップ461）。以上によりセクタディレクトリを復元することができる。以上のようにすれば、上位装置CPUより指示されたレコードの読み取りあるいは書き込み時に、ディレクトリを読み込めなくなるエラーが発生した場合であっても、チャンネルアダプタCAは自動的にディレクトリ復元データを用いてディレクトリを再構築するから、次のアクセスの際

(17)

特開平7-311661

31

して該ディレクトリを参照して高速のアクセスが可能となる。

【0075】図46は、ディレクトリにエラーが発生した場合の全体の概略シーケンスの説明図である。上位装置（ホスト）CPUよりレコード読み取り命令READ Dが発行されると、チャンネルアダプタCAは該コマンドを受け付け、ディレクトリの読み込み処理を行う。ディレクトリの読み込みエラーが発生すると、チャンネルアダプタCAはアクセス法を通常のディレクトリ使用のアクセス法からスキップサーチ法に切り替える。以後、図42に示す手順に従って目的とするレコードをスキップサーチし、目的レコードが求めればデータ部を上位装置に転送する。この場合、スキップサーチ処理と並行して図44及び図45のレコードディレクトリ、セクタディレクトリの再構成処理を行ってディレクトリを再構築することもできる。

【0076】(c-4) ディレクトリにエラーが発生した場合の別の処理

以上では、ディレクトリにエラーが発生した場合、自動的に目的のレコードをサーチして読み/書きし、またディレクトリを再構築した場合について説明した。しかし、かかる目的レコードのサーチやディレクトリの再構築は上位装置CPUからの指示に従って行うようにしても良い。図47はかかる場合のレコードリード手順説明図である。上位装置（ホスト）CPUよりレコード読み取り命令READ CKDが発行されると、チャンネルアダプタCAは該コマンドを受け付け、ディレクトリの読み込み処理を行う。この場合、ディレクトリの読み込みエラーが発生すると、チャンネルアダプタCAはエラー情報を作成して上位装置にエラー報告を行う。

【0077】上位装置CPUはエラーを解析し、特殊ダンプコマンドとレコード番号NをチャンネルアダプタCAに発行する。チャンネルアダプタCAは特殊ダンプコマンドを受け付け、図42に示す手順に従って目的とするレコードをサーチし、目的レコードが求めれば、カウント部、キー部、データ部を順次上位装置に転送する。以上は、特殊ダンプ命令による処理にディレクトリ再構築処理が伴わない場合である。ディレクトリ再構築処理が伴う場合には、図44及び図45のレコードディレクトリ、セクタディレクトリの再構築処理を行い、ディレクトリを再構築し、対象トラックの先頭に書き込む。

【0078】(c-5) コピー時にディレクトリエラーが発生した場合の処理

半導体ディスク装置にメモリ障害を生じると、サービスアダプタSAの制御によりエラーが生じた半導体メモリモジュールMSの内容を予備の半導体メモリモジュールHSにコピーし、しかる後、障害を生じた半導体メモリモジュールを新たな半導体メモリモジュールと交換する（第1実施例、図1参照）。かかる場合、エラーが発生したトラック以外のコピーはサービスアダプタSAが行

32

い、エラーが発生したトラックフィールドのコピーについてはチャンネルアダプタCAに依頼して該当チャンネルアダプタCAが行う。というのは、サービスアダプタSAはフィールド単位の処理ができず、固定長（64バイト）毎のブロックコピーをする。このため、エラーが発生したトラックフィールドのコピーについて、同様のコピー処理を行ってもエラー発生情報をディレクトリや各カウント部付加情報に書き込めず、ユーザデータのアクセスは保証できないものになってしまうからである。そこで、サービスアダプタSAはメモリ障害が発生したトラックフィールドのコピー処理についてはチャンネルアダプタCAに依頼し、依頼を受けたチャンネルアダプタCAは、ディレクトリから順にエンドマークEOFまでフィールド単位でコピー処理を行う。

【0079】チャンネルアダプタCAはコピーの途中でメモリ障害を起こしているフィールド（カウント部、キー部、データ部）を発見すると、そのフィールド（カウント部の場合にはレコード全体）のみを無効化し、カウント部における必要な制御情報の書き換えを行う。しかし、ディレクトリ部がメモリ障害を起こしている場合にディレクトリを無効化してしまうと、当該トラックフィールド上の全てのレコードが無効化してしまい、更にはチャンネルアダプタCAもディレクトリを読み込めないためトラックのフォーマットを判断できない。

【0080】これを回避するために、チャンネルアダプタCAは、固定位置のHAフィールド及びレコードROをコピーし、レコードR1以降のフィールドに関しては、スキップサーチ方法により次のレコードフィールドのアドレスを取得しながらエンドマークまでコピーを行う。すなわち、現レコードのカウント部を読み込めば、ディレクトリの再構築に必要な次のレコード迄のオフセットアドレスや現レコードのセクタ値、現レコードのキー長、データ長を取得でき、現レコードフィールドのコピーを行って次のレコードの処理に移れる。これをエンドマーク検出まで続けるとディレクトリを除く全てのフィールドのコピーができる。更に、上記コピー処理はディレクトリ再構築のために必要な全てのレコード部のディレクトリ復元データの読み込み処理を含んでいるため、必然的にディレクトリを再構築することができる。そこで、各フィールドのコピー処理をすると同時にディレクトリの再構築処理を行い、最後にチャンネルアダプタCA内部で再構築したディレクトリをコピー先のディレクトリ領域に書き込むことでメモリ交換処理時のユーザデータ消失を防止することができる。

【0081】図48及び図49は半導体メモリにエラーが生じた時のコピー処理のフロー図である。尚、半導体ディスク装置は図1に示す構成を有しているものとする。サービスアダプタSAはエラーが発生した半導体メモリモジュールMS及びトラックを認識すると（ステップ501）、該半導体メモリモジュールMSのコピーを

47

33

開始する(ステップ502)。コピー対象トラックがエラー発生トラックであるかチェックし(ステップ503)、エラー発生トラックでなければ該トラックのコピーを実行し(ステップ504)、コピー完了後半導体メモリモジュールMSのコピーが完了したチェックし(ステップ505)、完了していれば次のトラックについてステップ503以降の処理を行う。

【0082】一方、ステップ503において、コピー対象トラックがエラー発生トラックの場合には、該エラー発生トラックのコピーをチャンネルアダプタCAに依頼する(ステップ506)。チャンネルアダプタCAはフィールド単位でのコピーを開始し、フィールドの読み込みを行う(ステップ507)。該フィールドでエラーを検出すれば(ステップ508)、ディレクトリのエラーかチェックし(ステップ509)、ディレクトリのエラーでない場合には当該フィールドを無効化し(ステップ510)、ステップ507に飛ぶ。しかし、ディレクトリのエラーであればフラグFを"1"にし("1"→F、ステップ511)、ステップ507に飛ぶ。ステップ508において、フィールドにエラーが存在しない場合には、該フィールドのコピーを実行し(ステップ512)、ついで、F="1"かチェックし(ステップ513)、F="1"の場合にはディレクトリの再構築処理を行う(ステップ514)。

【0083】しかる後、エンドマークEOFが検出されたチェックし(ステップ515)、検出してなければステップ507に飛び以降のフィールドのコピーを実行する。エンドマークEOFが検出された場合には、F="1"かチェックし(ステップ516)、F="0"の場合にはサービスアダプタSAにコピー終了を通知する(ステップ517)。これにより、サービスアダプタは次のトラックに対してステップ503以降のコピー処理を繰り返す。ステップ516において、F="1"の場合には、再構築したディレクトリをコピー先のディレクトリ領域に書き込み(ステップ518)、ついでFを"0"にし("0"→F、ステップ519)、サービスアダプタSAにコピー終了を通知する(ステップ517)。これにより、サービスアダプタは次のトラックに対してステップ503以降のコピー処理を繰り返す。

【0084】以上のように、第3実施例によれば、ディレクトリを読み込めなくなっても、該ディレクトリを用いて各レコードに含まれる制御情報復元データを用いて上位装置により指示された目的レコードへのアクセスが可能となる。又、制御情報復元データを用いてディレクトリの再構築ができるため、以後、再構築したディレクトリを用いてアクセスができる。このため、アクセス速度が低下することはない。又、メモリエラーによる半導体メモリモジュール交換時のコピー処理において、ディレクトリを読み込めなくなるエラーが検出された場合にも、各レコードに含まれる制御情報復元データを用い

(18)

特開平7-311661

34

て順次次のレコード位置が判明するためレコードのコピーができ、しかも、ディレクトリを再構築してコピー先のディレクトリ領域に書き込むことができる。

【0085】(d) 第4実施例

(d-1) 全体の構成

図50はデータを圧縮して半導体ディスクに記憶する半導体ディスク装置の実施例の構成図であり、二重化構成になっており、添字1を有するモジュールは第1の半導体ディスク装置G0側のモジュールであり、添字2を有するモジュールは第2半導体ディスク装置G1側のモジュール、添字の無いモジュールは共通のモジュールである。

【0086】CA₁、CA₂は上位装置のCPU、チャンネルとのインターフェース制御を行うチャンネルアダプタであり、上位装置(チャンネル)に接続されるようになっている。RM₁、RM₂はリソースマネージャであり排他制御等の処理動作を制御すると共にサブシステム全体の資源管理を行う。TS₁、TS₂は排他制御テーブル等のテーブル類を記憶するテーブルストロージ(テーブル記憶部)、SA₁、SA₂はサービスアダプタ、MDKは半導体ディスクの内容を一時的にバックアップする磁気ディスク装置、DA₁、DA₂は磁気ディスク装置とのインターフェース制御を行うデバイスアダプタ、BANK₁、BANK₂は半導体ディスクであり、複数の半導体メモリモジュール及び予備の半導体メモリモジュールが搭載されるようになっている。ESP₁、ESP₂は半導体ディスクへのアクセス制御を行うポート、MT₁、MT₂は半導体メモリを多数の所定サイズのブロックに分割した時、各ブロックの使用/未使用状態を記憶する半導体メモリテーブル、MCA₁、MCA₂は半導体メモリにおけるブロック毎の使用/未使用状態を管理し、チャンネルアダプタCA₁、CA₂からのブロック要求により未使用ブロックを割り当て、該ブロックを使用中にすると共に、チャンネルアダプタからのブロック解放要求により所定の使用中ブロックを未使用中に変更する。

【0087】BUS₁、BUS₂はバスであり、制御用バス、データ転送用バス、サービスバスを有している。CBUSは一方のテーブル記憶部TS₁(TS₂)の内容が変化した時、直ちに他方のテーブル記憶部TS₂(TS₁)にコピーするコピーバス、CBUS'は一方の半導体メモリテーブルMT₁(MT₂)の内容が変化した時、直ちに他方の半導体メモリテーブルMT₂(MT₁)にコピーするコピーバスである。第1、第2半導体ディスク装置G0、G1は中央点線を中心に対称に構成されており、上位CPUは対称に第1、第2半導体ディスク装置のチャンネルアダプタCA₁、CA₂に接続されている。

【0088】(d-2) 半導体メモリの構成

半導体メモリBANKは図51に示すように、トラックエミュレーションエリア51と、レコードのうちデータ

(10)

特開平7-311661

35

部を圧縮して記憶する圧縮データ格納エリア52に分けられている。圧縮データ格納エリア52は所定サイズの多数のブロック(論理ブロック)に分割され、各レコードのデータ部は圧縮された後、必要数のブロックに連続的にあるいは分散して格納される。

【0089】トラックエミュレーションエリア51の各トラックはレコードのデータ部が記憶されない点を除けば、図36に示すトラックフォーマットを備えている。各トラックフィールドには図52に示すように、ディレクトリ、ホームアドレスHA、レコード0のカウント部及びデータ部、各レコードのカウント部、キー部が記憶される。各レコードRi(i=1,2,...)のカウント部Cにはフィールドを制御する制御情報やレコード番号CCHHR、キー長KLデータ長DL等が書き入れ、更に、データ部が記憶される圧縮データ格納エリア52のブロック番号(ブロックアドレス)が書かれるようになっている。レコードを書き込む、図53に示すように書き込むべきレコードのデータ部DTを圧縮して圧縮データDT'とし、ついで、該圧縮データDT'のサイズに応じた数のブロック(図では3つ)確保し、該確保したブロックB1~B3に圧縮データDT'の第1~第3部分①~③を格納し、これらブロックB1~B3のブロックアドレスをカウント部に挿入してトラックエミュレーションエリア51に格納する。

【0090】(d-3) チャンネルアダプタの構成

図54は圧縮/復号処理を行うチャンネルアダプタの構成図である。チャンネルアダプタ内部において、61は非圧縮データを記憶する大容量(例えば1トラック分)の第1データバッファ、62は圧縮データを記憶する大容量(例えば1トラック分)の第2のデータバッファ、63はデータの圧縮・復号を行うデータ圧縮機構部、64は圧縮・復号動作を補助する辞書メモリ、64はチャンネルインターフェースプロトコル制御部、65は第1、第2のデータバッファからのデータを選択して出力すると共に、逆にデータを第1、第2のデータバッファに選択的に出力するセクタ、66は半導体ディスク装置内部の他のユニットと接続された内部バスを制御する内部バスインターフェース制御部(BIL)、67は非圧縮データのバイト数を計数する第1のカウント、68は圧縮データのバイト数を計数する第2のカウント、69はこれらハードウェア資源をマイクロプログラムにより制御するMPU、70はプログラム、制御テーブルその他のテーブル類を格納するコントロールストレージ、71はバスである。

【0091】(d-4) レコード読み取り処理

図55はレコード読み取り処理のシーケンス説明図である。上位装置CPUよりシークコマンドSKが発生すると、チャンネルアダプタCA(MPU69)は直ちに動作終了信号を上位装置に返す。上位装置CPUは動作終了信号を受信すれば、つぎにセットセクタコマンドSSを

36

発行する。以後、同様にして上位装置はサーチIDコマンドSID及びリードコマンド(READ CKD)を発行し、チャンネルアダプタCAはこれらコマンドを受信する。チャンネルアダプタCAのMPU69はリードコマンドを受信すると、制御テーブルを用いてシークコマンドで指示されている論理アドレス(トラックアドレスCCHH)を半導体ディスクの物理アドレスに変換する。この物理アドレスはトラックエミュレーションエリア51における対応トラックの先頭アドレス、換言すればディレクトリの先頭アドレスを示しているから、半導体ディスクよりディレクトリを読み込む。

【0092】ついで、セットセクタコマンドで指示されているセクタ値に応じたレコード番号をセクタディレクトリより求める(セットセクタ動作)。しかる後、該レコード番号のアドレスをレコードディレクトリより求め、該アドレスよりレコードのカウント部を読み取り、該カウント部に含まれるレコードのCCHHRとサーチIDコマンドで指示されている目的レコードのCCHHRが一致するかチェックする。一致すればサーチID動作を終了する。一致しない場合には次のレコード番号のレコードを読み取って同様の処理を行う。目的のレコードが見つければ、内部バスインターフェース制御部(BIL)66を介して該レコードのカウント部C、キー部Kを読み出す。セクタ65は該カウント部C、キー部Kを第1のデータバッファ61にセットし、チャンネルインターフェースプロトコル制御部64は該カウント部、キー部を上位装置に転送する。ついで、チャンネルアダプタCAのMPUはカウント部Cに書き込まれているブロックアドレス(1つに限らない)をBILにセットする。これにより、以後BILはブロックアドレスが指示するブロックよりデータを順次読み取りセクタ65を介して第2のデータバッファ62に書き込む。データ圧縮機構63は圧縮データを復元し、第1データバッファ61にセットし、チャンネルインターフェースプロトコル制御部64は該データを上位装置に転送する。

【0093】(d-5) レコード書替え処理

図56はレコード書替え処理のフロー図である。ライトコマンド(WRITE D)を受付けるまではリードコマンドの処理と同様である。チャンネルインターフェースプロトコル制御部64は上位装置からライトコマンド(WRITE D)を受信するとレコードを一時的に大容量のデータバッファ61に格納する(ステップ601)。ついで、チャンネルアダプタCAのMPU69は制御テーブルを用いてシークコマンドで指示されている論理アドレス(トラックアドレスCCHH)を半導体ディスクの物理アドレスに変換する。この物理アドレスは対応トラックのディレクトリ位置を示しているから、該ディレクトリを読み込む(ステップ602)。

【0094】しかる後、リードコマンドの場合と同様に目的レコードのアドレスを求め、該アドレスより目的レ

37

コードのカウンタ部を読み取る(ステップ603)。そして、カウンタ部に含まれるブロックアドレス及びブロック数 B_0 を保存する(ステップ604)。尚、このブロックアドレス及びブロック数は、目的レコードのデータ部が記憶されているブロックのアドレス及びその数を示す。ブロックアドレス及びブロック数 B_0 の保存が完了すれば、あるいは以上と並行してデータ圧縮機構63はレコードのうちデータ部を圧縮して第2のデータバッファ62に格納する(ステップ605)。圧縮データのバイト数は第2のカウナ68が監視しているから、該バイト数を1ブロックの容量で除算することにより必要なブロック数 B_n を求める(ステップ606)。ブロック数 B_n が求めれば、今までのブロック数 B_0 との大小を比較し(ステップ607)、 $B_n = B_0$ であれば、BILに今までのブロックアドレスをセットし(ステップ608)、該ブロックアドレスが指示するブロックに圧縮データを書き込む(ステップ609)。

【0095】一方、 $B_n < B_0$ の場合には、MPU69は不要になった $(B_0 - B_n)$ 個のブロックアドレスを半導体メモリ管理アダプタMCTに指示する。これにより、半導体メモリ管理アダプタMCTは指示されたブロックアドレスを未使用にするようにメモリテーブルMTの内容を変更する(ステップ610)。又、MPU69は目的レコードのカウンタ部より、解放指示したブロックのブロックアドレスを削除する(ステップ611)。以後、チャンネルアダプタのMPU69はBIL66に必要な数のブロックアドレスをセットし(ステップ608)、該ブロックアドレスが指示するブロックに圧縮データを書き込む(ステップ609)。

【0096】又、 $B_n > B_0$ の場合には、MPU69は $(B_n - B_0)$ 個の不足ブロックの割り当てを半導体メモリ管理アダプタMCTに要求する(ステップ612)。半導体メモリ管理アダプタMCTは該ブロック要求を受信すると、メモリテーブルMTを参照して要求された数の未使用ブロックを求めて使用中に変更すると共に、該求めた未使用ブロックのブロックアドレスをチャンネルアダプタCAに通知する。チャンネルアダプタCAのMPU69は半導体メモリ管理アダプタMCTより必要な数のブロックアドレスを受信すれば(ステップ613)、目的レコードのカウンタ部に新たに割り当てられたブロックのブロックアドレスを追加する(ステップ614)。以後、チャンネルアダプタのMPU69はBIL66に新レコードの書き込みに必要な数のブロックアドレスをセットし(ステップ608)、該ブロックアドレスが指示するブロックに順次圧縮データを書き込む(ステップ609)。

【0097】以上、第4の実施例によれば、ユーザデータを圧縮して記憶するからメモリの有効利用が図れ、しかも、制御情報(ディレクトリ、カウンタ部等)は圧縮しないため、復元処理が不要となり上位装置のデータア

(20)

特開平7-311661

38

クセス時間の短縮を図ることができる。又、第4実施例によれば、データを圧縮して半導体メモリに書き込む場合、効果的にメモリ領域の解放、割当てを行ってメモリの有効利用を図ることができる。以上、本発明を実施例により説明したが、本発明は請求の範囲に記載した本発明の主旨に従い種々の変形が可能であり、本発明はこれらを排除するものではない。

【0098】

【発明の効果】

10 第1発明(第1実施例)の効果

以上第1の発明によれば、予備の半導体メモリモジュールを追加するだけで、メモリエラーが生じた半導体メモリモジュールの記憶内容を無停止で退避することができ、しかも、大がかりな保守装置を必要としない。また、第1発明によれば、半導体メモリモジュールをアクセス制御単位に細分化し、該アクセス制御単位でコピーを行う。このため、該制御単位で排他制御を行うことができ、コピー中であっても上位装置はコピー対象でない他のアクセス制御単位部分にアクセスすることができ、上位装置のアクセスに悪影響を与えることがない。更に、第1発明によれば制御単位のコピーが完了する毎にサービスアダプタSAは、チャンネルアダプタCA等に設けられている制御テーブルCTLにおける物理アドレスをコピー先の物理アドレスに変更する。このため、アクセス制御単位のコピー完了後、該アクセス制御単位については直ちにコピー先(予備の半導体メモリモジュール)にアクセスすることができる。

【0099】また、第1発明によれば、アクセス制御単位毎にコピーが正常に完了したか否かを管理するコピー管理テーブルを設け、チャンネルアダプタCA等のモジュールはデータリード時に該コピー管理テーブルを参照し、アクセス対象である領域(トラック)が正常の場合にはアクセスし、異常の場合にはアクセスせずエラーとする。この結果、コピー時にコピーエラーが発生しても誤ったデータを読み取って処理することがなく、誤動作を防止できる。更に第1発明によれば、チャンネルアダプタ等のモジュールはコピー中の半導体メモリモジュールにデータを書き込む際、フォーキング処理に従ってオリジナルな物理アドレスとコピー先の物理アドレスが示す2つの記憶エリアにデータを書き込む。この結果、制御単位のコピー完了毎に制御テーブルを変更する必要がなく、コピーを高速に行うことができる。又、第1発明によれば、コピーを停止して上位装置CPUの書き込みコマンドを実行するようにしたから上位装置のアクセス速度は低下しない。又、この場合、アクセス単位がコピー単位より大きいければ、アクセス単位部分のコピーが完了したものとみなすからコピー速度を向上することができる。

【0100】第2の発明の効果、

50 第2の発明によれば、半導体ディスクの構成を変更して

39

も、構成変更前にバックアップディスク装置に退避したデータを構成変更後の半導体ディスク装置に正しく復元でき、変更前のユーザデータが無効化されることがなく使用することができる。従って、半導体ディスクの構成を適宜に自由に変更することができる。

・第3の発明

第3の発明によれば、各レコードのカウント部にディレクトリ復元データ（制御情報復元データ）を含ませたから、エラーによりディレクトリを読み込めなくなっても、各レコードに含まれる制御情報復元データを用いて、上位装置から指示された目的レコードへのアクセスが可能となる。又、制御情報復元データを用いてディレクトリの再構築ができたため、以後、再構築したディレクトリを用いてアクセスができる。このため、上位装置のアクセス速度は低下することがない。又、第3発明によれば、エラーによる半導体メモリモジュール交換時のコピー処理において、ディレクトリでエラーが検出された場合にも、各レコードに含まれる制御情報復元データを用いて順次、次のレコード位置が判明するためレコードのコピーができ、しかも、ディレクトリを再構築してコピー先のディレクトリ領域に書き込むことができる。

【0101】・第4の発明の効果

第4の発明によれば、ユーザデータを圧縮して記憶するからメモリの有効利用が図れ、しかも、制御情報（ディレクトリ、カウント部等）は圧縮しないため、復元処理が不要となり上位装置のデータアクセス時間の短縮を図ることができる。又、データを圧縮して半導体メモリに書き込む場合、効果的にメモリ領域の解放、割当て制御を行うからメモリの有効利用を図ることができる。

【図面の簡単な説明】

【図1】本発明の半導体ディスク装置の構成図（第1実施例）である。

【図2】制御テーブルの説明図である。

【図3】コピー管理テーブルの説明図である。

【図4】半導体ディスクの構成図である。

【図5】制御テーブルの内容（初期状態）説明図である。

【図6】コピー処理説明図である。

【図7】制御テーブルの内容（中間）説明図である。

【図8】制御テーブルの内容（コピー完了）説明図である。

【図9】コピー処理のフロー図（その1）である。

【図10】コピー処理のフロー図（その2）である。

【図11】コピー管理テーブルの説明図である。

【図12】半導体メモリモジュールの交換後の処理フロー図である。

【図13】半導体メモリモジュールの交換後の別の処理フロー図である。

【図14】コピー管理テーブルを用いたアクセス処理フロー図である。

(21)

特開平7-311661

40

【図15】制御テーブルの別の構成図である。

【図16】制御テーブルの内容（初期時）説明図である。

【図17】制御テーブルの内容（変更指示後）説明図である。

【図18】フォーキング処理説明図である。

【図19】コピー処理説明図である。

【図20】制御テーブルの内容（物理アドレス交換指示後）説明図である。

10 【図21】制御テーブルの内容（コピー終了時）説明図である。

【図22】コピー処理のフロー図（その1）である。

【図23】コピー処理のフロー図（その2）である。

【図24】コピー中の制御単位エリアにライト命令が発生した場合の処理フロー図である。

【図25】半導体ディスク装置の実際の構成図である。

【図26】各モジュールの構成図である。

【図27】本発明の半導体ディスク装置の別の構成図（第2実施例）である。

20 【図28】論理ドライブの構成図である。

【図29】論理ドライブ構成の変更説明図である。

【図30】論理ドライブ変更時の従来の問題点説明図である。

【図31】論理ドライブ構成変更の別の説明図である。

【図32】本発明の復元処理の概略説明図である。

【図33】アドレス変換テーブル説明図である。

【図34】ユーザデータ復元説明図である。

【図35】本発明の復元処理のフロー図である。

30 【図36】半導体ディスクにおけるトラックフォーマット説明図である。

【図37】ディレクトリの構成図である。

【図38】カウント部の構成図である。

【図39】トラックフィールドの構成説明図である。

【図40】レコードの読み取りシーケンス説明図である。

【図41】レコード書き込みシーケンス説明図である。

【図42】カウント部付加情報を用いたスキップサーチ処理のフロー図である。

40 【図43】カウント部付加情報を用いたスキップサーチ処理の説明図である。

【図44】レコードディレクトリの再構成処理のフロー図である。

【図45】セクタディレクトリの再構成処理のフロー図である。

【図46】ディレクトリにエラーが発生した場合の全体のシーケンス説明図である。

【図47】トラック特殊アクセス法を用いたホストアクセスの説明図である。

50 【図48】半導体メモリにエラーが発生した時のコピー処理フロー図（その1）である。

(22)

特開平7-311661

41

【図49】半導体メモリにエラーが発生した時のコピー処理フロー図(その2)である。

【図50】半導体ディスク装置の構成図(第4実施例)である。

【図51】メモリマップ説明図である。

【図52】トラックエミュレーションエリアの説明図である。

【図53】ユーザデータの論理ブロックへの格納説明図である。

【図54】チャネルアダプタの構成図である。

【図55】レコードの読み取りシーケンス説明図である。

42

【図56】レコードの書替え処理のフロー図である。

【図57】従来の半導体ディスク装置の構成図である。

【符号の説明】

11a-1, 11a-2・・・上位装置

12・・・半導体ディスク装置

13・・・半導体ディスク制御装置

13a-1, 13a-2・・・チャネルアダプタ

13b・・・リソースマネージャ

13c・・・サービスアダプタ

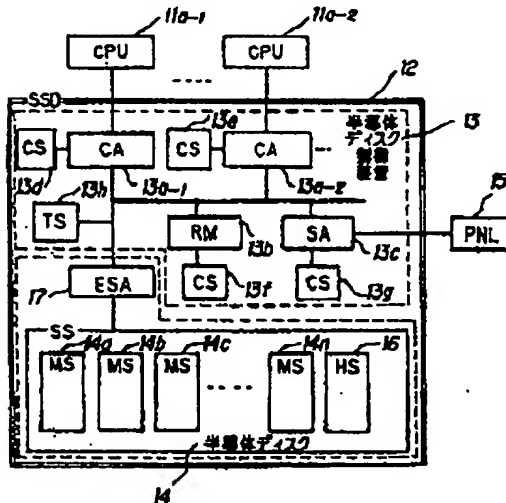
10 14・・・半導体ディスク

14a～14n・・・複数の半導体メモリモジュール

16・・・予備の半導体メモリモジュール

【図1】

本発明の半導体ディスク装置の構成図
(第1実施例)



【図2】

制御テーブルの説明図

項目	説明
1	物理アドレス
2	物理アドレス
...	...
n	物理アドレス

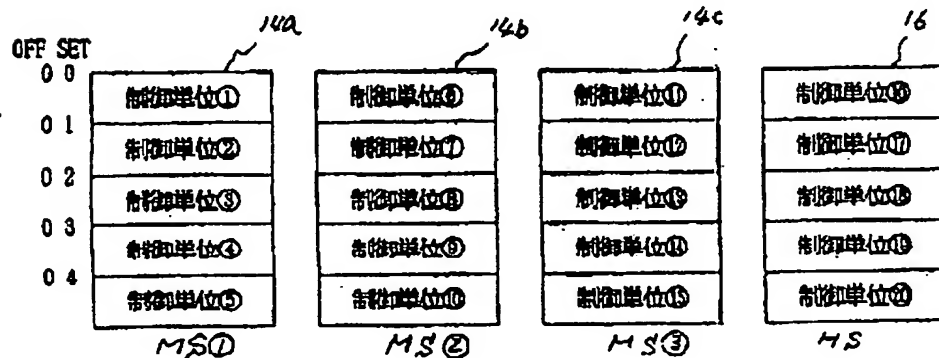
【図3】

コピー管理テーブルの説明図

表		有効/無効
シリンダ	トラック	

【図4】

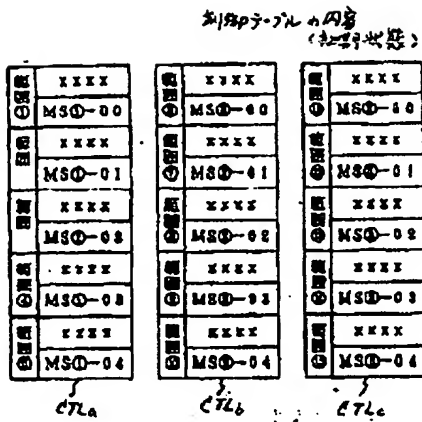
半導体ディスクの構成



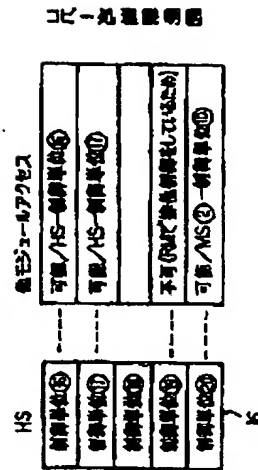
(23)

特開平7-311661

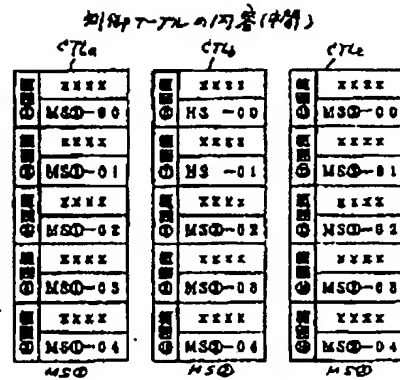
【図5】



【図6】

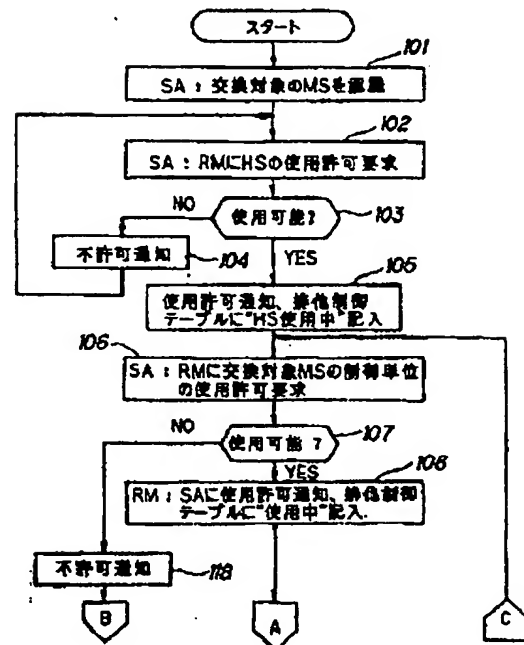


【図7】

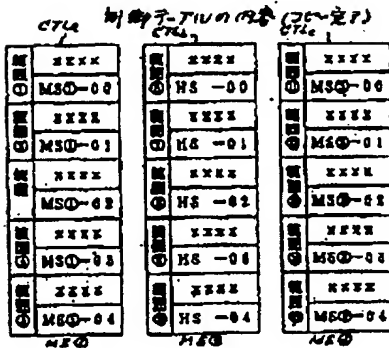


【図9】

コピー処理のフロー図(その1)

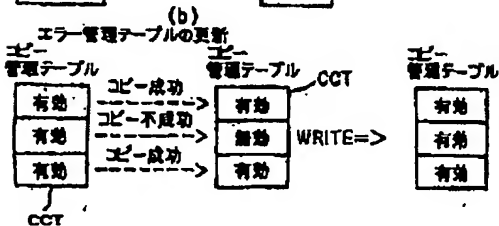
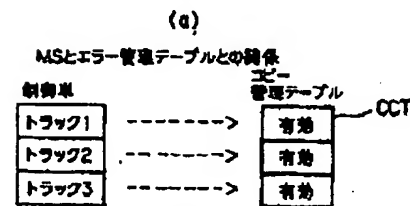


【図8】

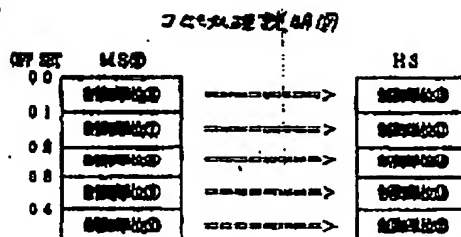


【図11】

コピー管理テーブルの説明図



【図19】

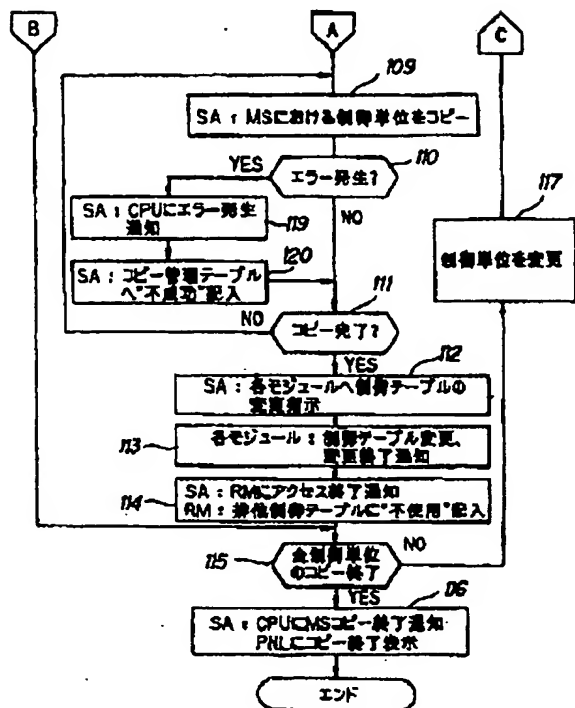


(24)

特開平7-311661

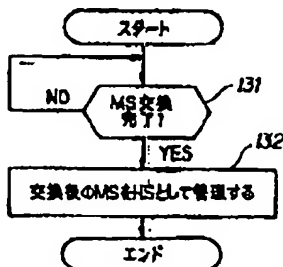
【図10】

コピー処理のフロー図(その2)



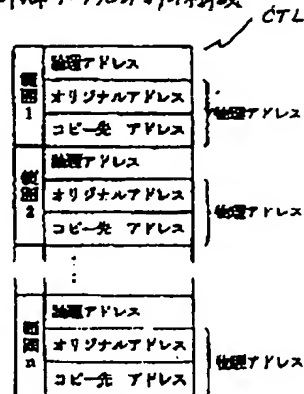
【図12】

半導体メモリモジュール交換後の処理フロー



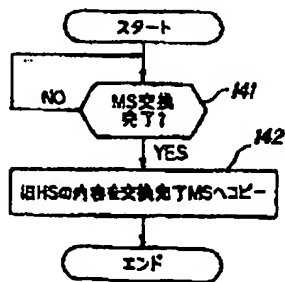
【図15】

制御テーブルの別構成



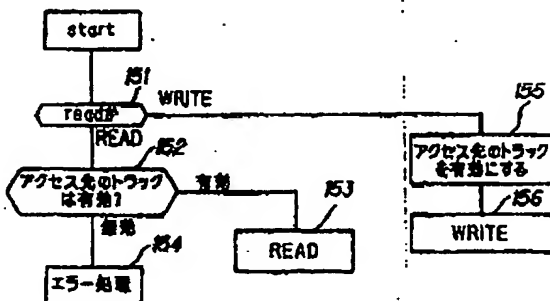
【図13】

半導体メモリモジュール交換後の別の処理フロー



【図14】

コピー管理テーブルを用いたアクセス処理のフロー図



(25)

特開平7-311661

【図16】

制御テーブルの内容(初期時)

①図面	XXXX	①図面	XXXX	①図面	XXXX
MSD-00	無効	MSD-00	無効	MSD-00	無効
②図面	XXXX	②図面	XXXX	②図面	XXXX
MSD-01	無効	MSD-01	無効	MSD-01	無効
③図面	XXXX	③図面	XXXX	③図面	XXXX
MSD-02	無効	MSD-02	無効	MSD-02	無効
④図面	XXXX	④図面	XXXX	④図面	XXXX
MSD-03	無効	MSD-03	無効	MSD-03	無効
⑤図面	XXXX	⑤図面	XXXX	⑤図面	XXXX
MSD-04	無効	MSD-04	無効	MSD-04	無効

CTL_a CTL_b CTL_c

【図17】

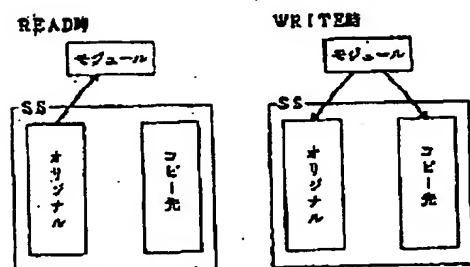
制御テーブルの内容(変更指示後)

①図面	XXXX	①図面	XXXX	①図面	XXXX
MSD-00	無効	MSD-00	HS-00	MSD-00	無効
②図面	XXXX	②図面	XXXX	②図面	XXXX
MSD-01	無効	MSD-01	HS-01	MSD-01	無効
③図面	XXXX	③図面	XXXX	③図面	XXXX
MSD-02	無効	MSD-02	HS-02	MSD-02	無効
④図面	XXXX	④図面	XXXX	④図面	XXXX
MSD-03	無効	MSD-03	HS-03	MSD-03	無効
⑤図面	XXXX	⑤図面	XXXX	⑤図面	XXXX
MSD-04	無効	MSD-04	HS-04	MSD-04	無効

CTL_a CTL_b CTL_c

【図18】

フォーマット説明図



【図20】

制御テーブルの内容(物理作戻指示後)

①図面	XXXX	①図面	XXXX	①図面	XXXX
MSD-00	無効	MSD-00	HS-00	MSD-00	無効
②図面	XXXX	②図面	XXXX	②図面	XXXX
MSD-01	無効	MSD-01	HS-01	MSD-01	無効
③図面	XXXX	③図面	XXXX	③図面	XXXX
MSD-02	無効	MSD-02	HS-02	MSD-02	無効
④図面	XXXX	④図面	XXXX	④図面	XXXX
MSD-03	無効	MSD-03	HS-03	MSD-03	無効
⑤図面	XXXX	⑤図面	XXXX	⑤図面	XXXX
MSD-04	無効	MSD-04	HS-04	MSD-04	無効

【図33】

アドレス変換テーブル

読取アドレス0	読取アドレス1	読取アドレス2	読取アドレス3	読取アドレス4
00000000	00000001	00000002	00000003	00000004
読取アドレス数 20	読取アドレス数 10	読取アドレス数 10	読取アドレス数 10	読取アドレス数 10
読取アドレス0	読取アドレス10	読取アドレス20	読取アドレス30	読取アドレス40
読取アドレス数 10	読取アドレス数 10	読取アドレス数 10	読取アドレス数 10	読取アドレス数 10

特開平 7 - 3 1 1 6 6 1

【例 24】

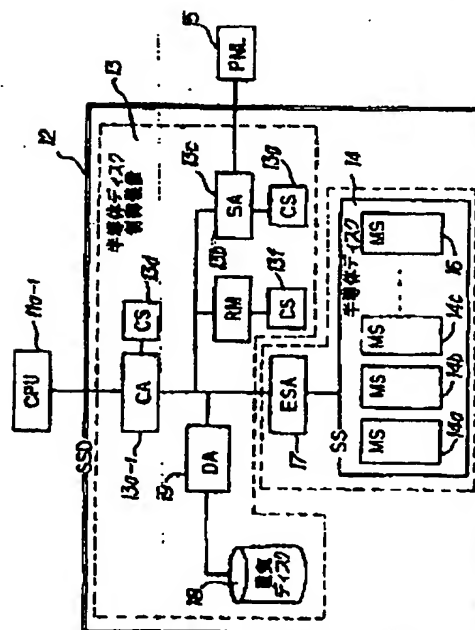
ブロー中の制御単位エリアでWRITE命令
が発生した場合の処理フロー

```

graph TD
    Start([スタート]) --> 251[CPUはVCAにWRITEコマンド]
    251 --> 252[CA: RMにアクセス許可要求]
    252 --> 253{コピー中か?}
    253 -- NO --> 254[アクセス許可]
    253 -- YES --> 255[RM: SAへWRITEコマンド発生を通知]
    255 --> 256[SA: コピー停止]
    256 --> 257[RM: アクセス許可]
    257 --> 258{書き込み完了?}
    258 -- NO --> 259{SA/CAの  
アクセス単位は  
コピー単位?}
    258 -- YES --> 259
    259 -- NO --> 261[当該コピー単位を再度コピー]
    259 -- YES --> 260[当該アクセス単位に  
対してコピー完了とみなす]
    261 --> 260
    260 --> End([終了])
  
```

【圖 27】

本発明の半導体ディスク装置の別の構成例
(第2実施例)



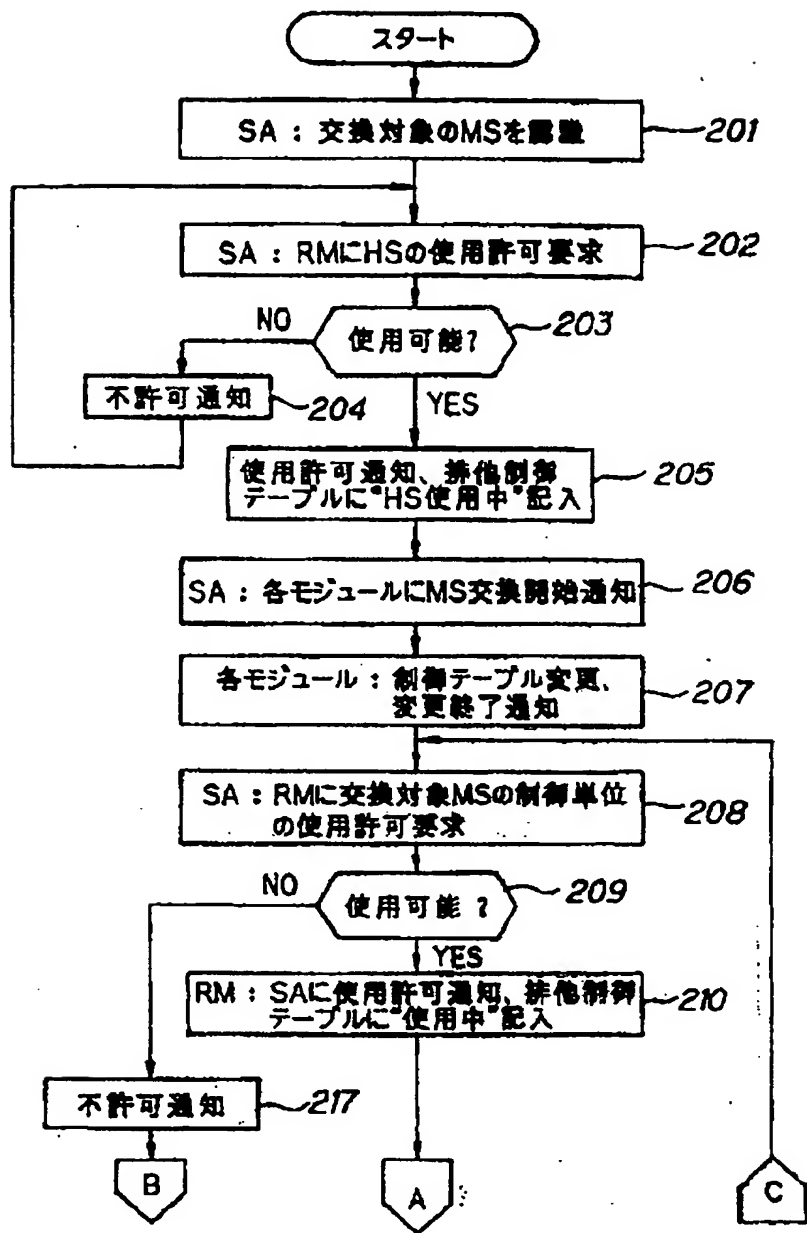
55

(27)

特開平7-311661

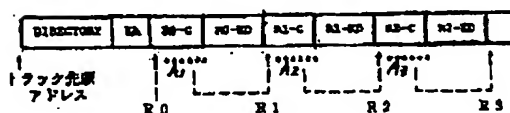
【図22】

コピー処理のフロー (その1)



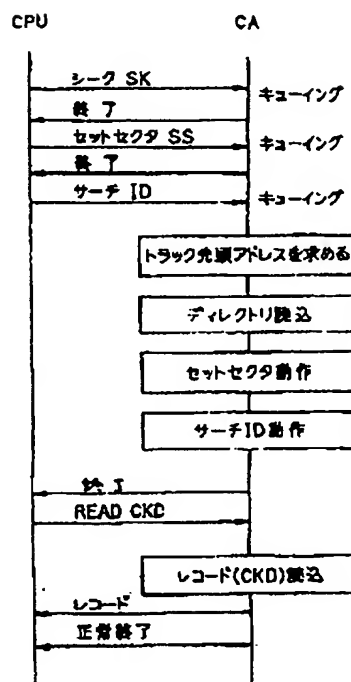
【図43】

カウント部付加情報を用いたスキップサーチ処理



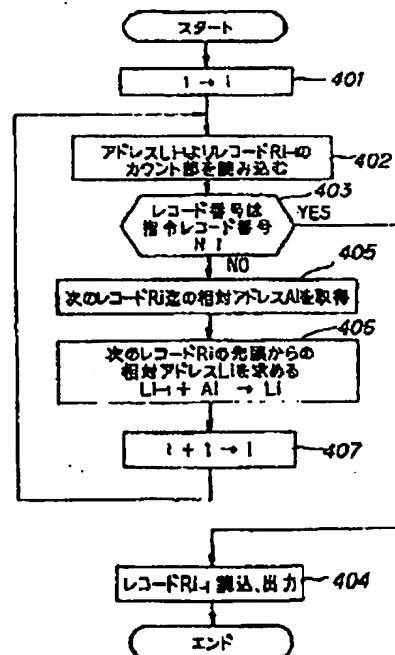
【図40】

レコード読取りシーケンス



【図42】

カウント部付加情報を用いたスキップサーチ処理

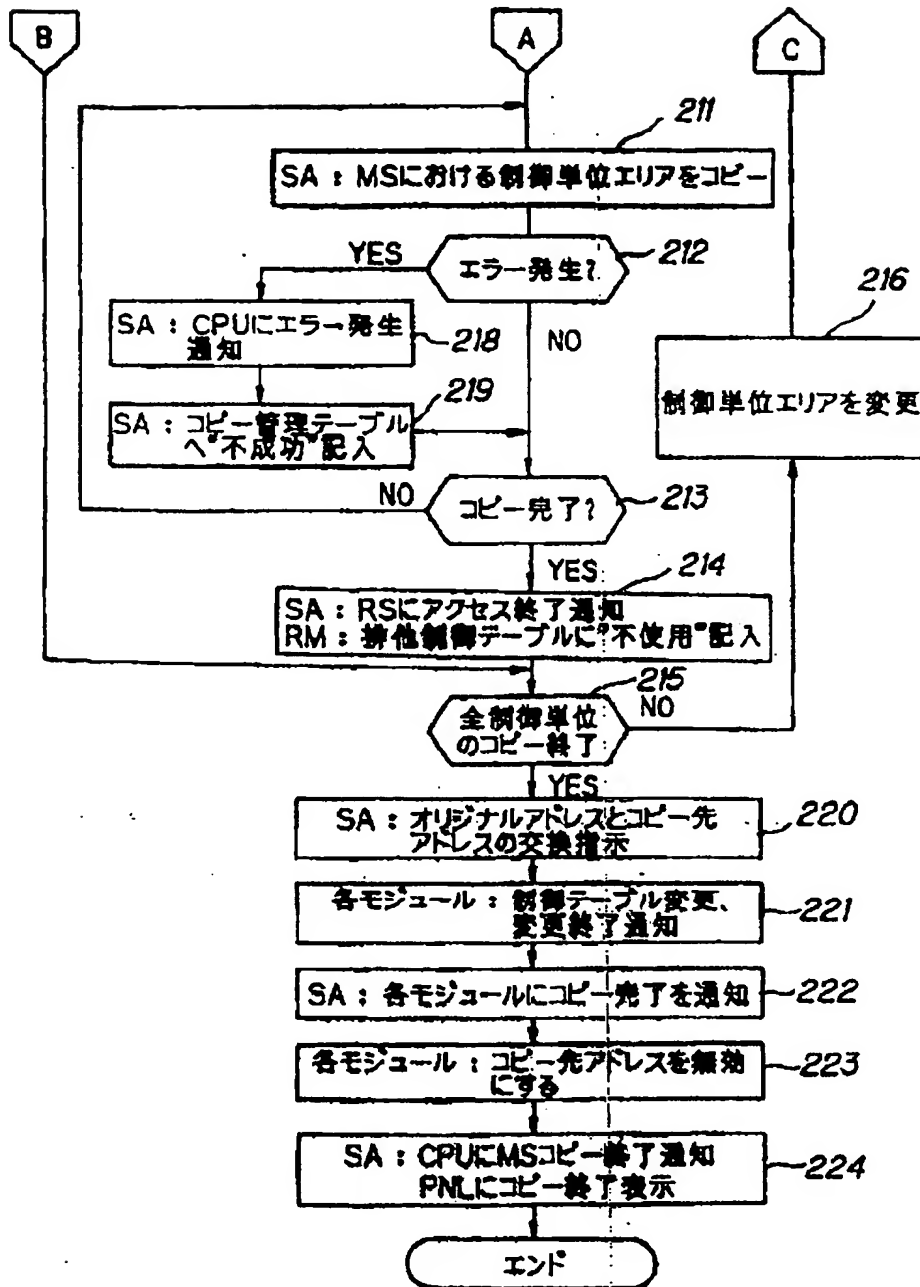


(28)

特開平7-311661

【図23】

コピー処理のフロー（その2）

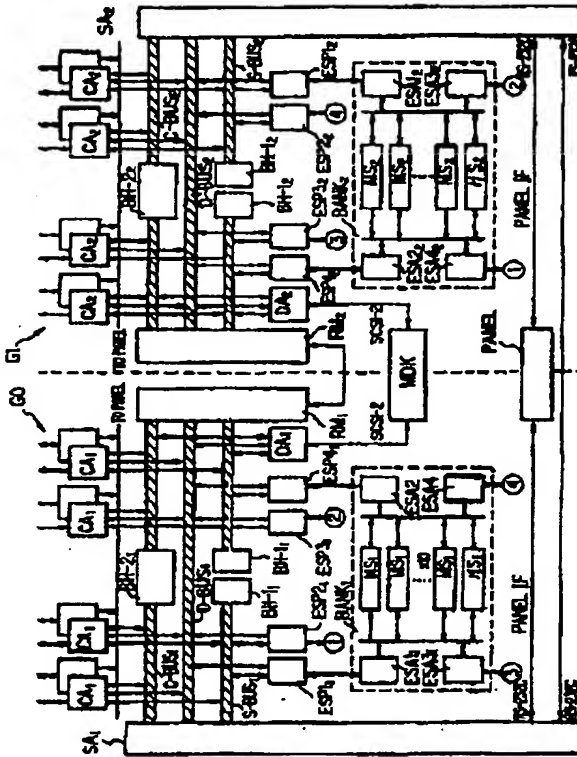


(29)

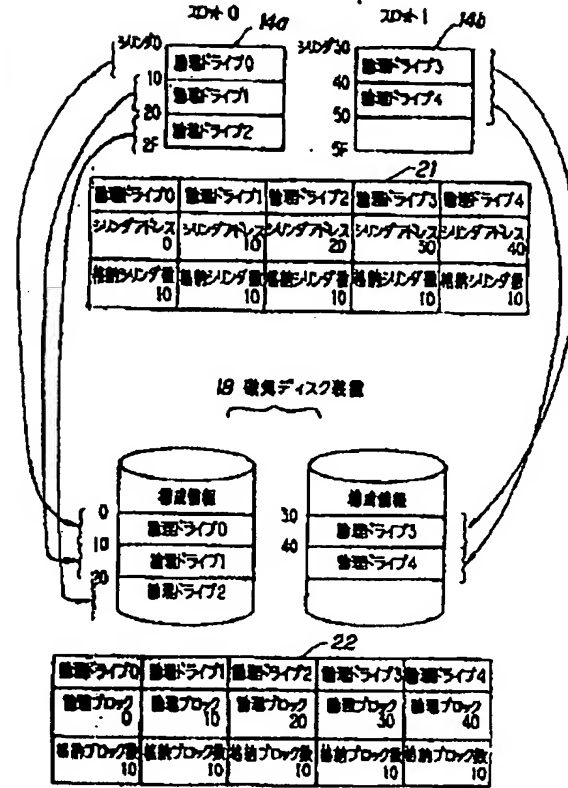
特開平7-311661

【図25】

半導体ディスク装置の内部の構成図



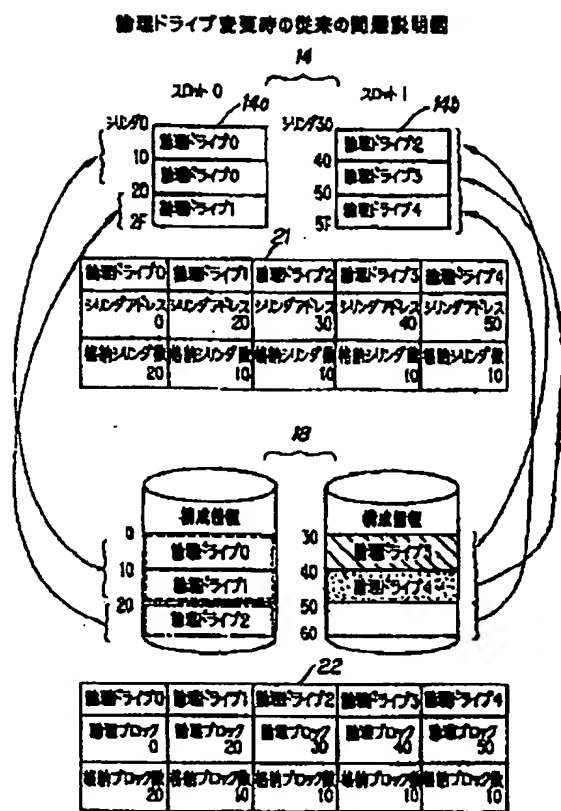
【図28】

論理ドライブの構成
14 半導体ディスク

(30)

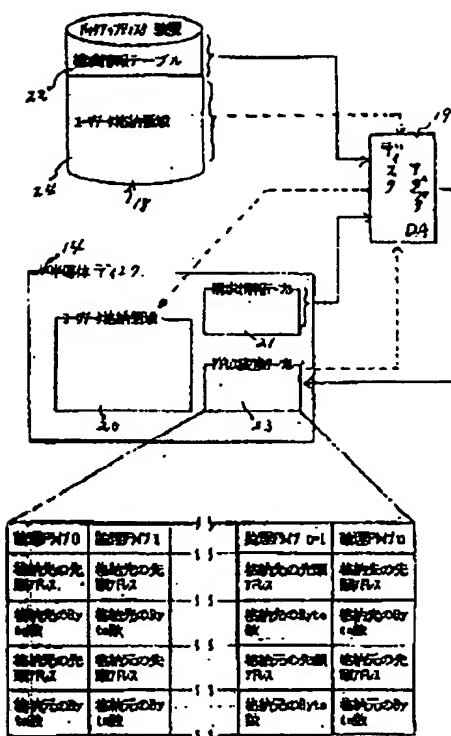
特開平7-311661

【図30】

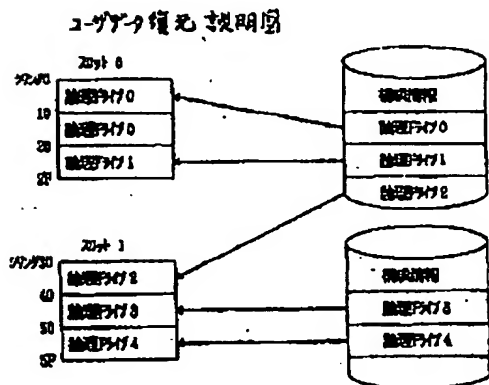


【図32】

本発明の援元処理の概略説明図



【図34】



【図36】

半導体ディスクにフォーマット

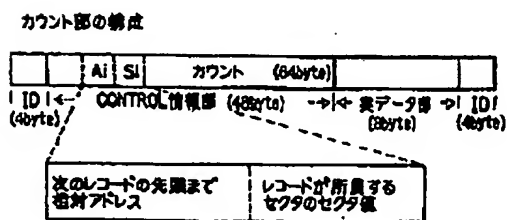
形成情報により定義された固定メモリアドレス

TRACK 0	DIRECTORY	RA	RA-C	RA-D	RA-E	RA-F	RA-G
TRACK 1	DIRECTORY	RA	RA-C	RA-D	RA-E	RA-F	RA-G
TRACK 2	DIRECTORY	RA	RA-C	RA-D	RA-E	RA-F	RA-G

(← 固定長 →)

【図51】

【図38】



フォーマット構成説明図

Track Emulation

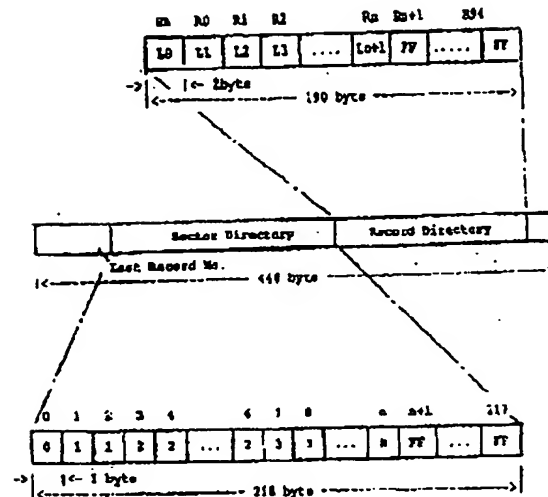
Track Emulation	RA	RA-C	RA-D	RA-E	RA-F	RA-G	RA-H
Track 0							
Track 1							
Track 2							

Logical Block

特開平 7 - 3 1 1 6 6 1

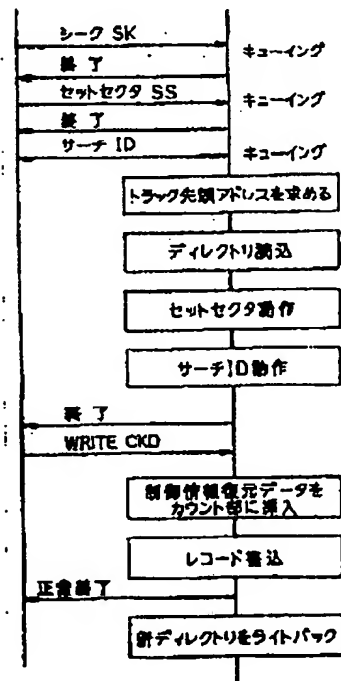
【圖 37】

ニルサリの構成図



レコード音読みシーケンス

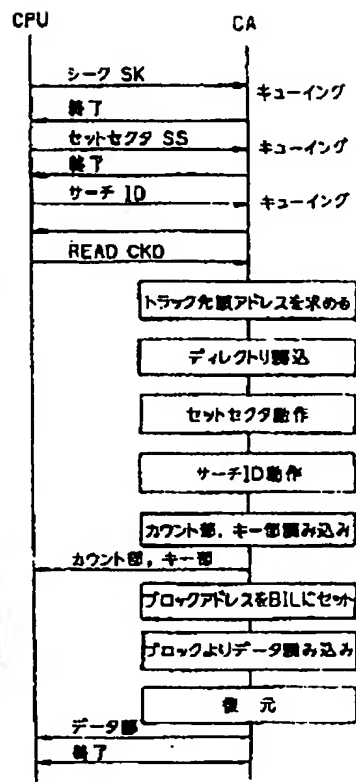
トウチノ・ワールドの構成・説明図



特開平 7 3 1 1 6 6 1

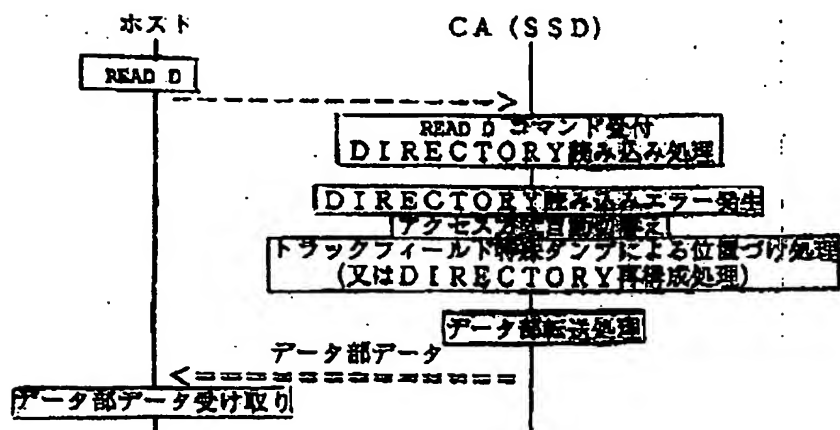
【例 55】

レコード屋取りシーケンス



【例 46】

ディレクトリにエラが発生した場合の
全体のシナシス

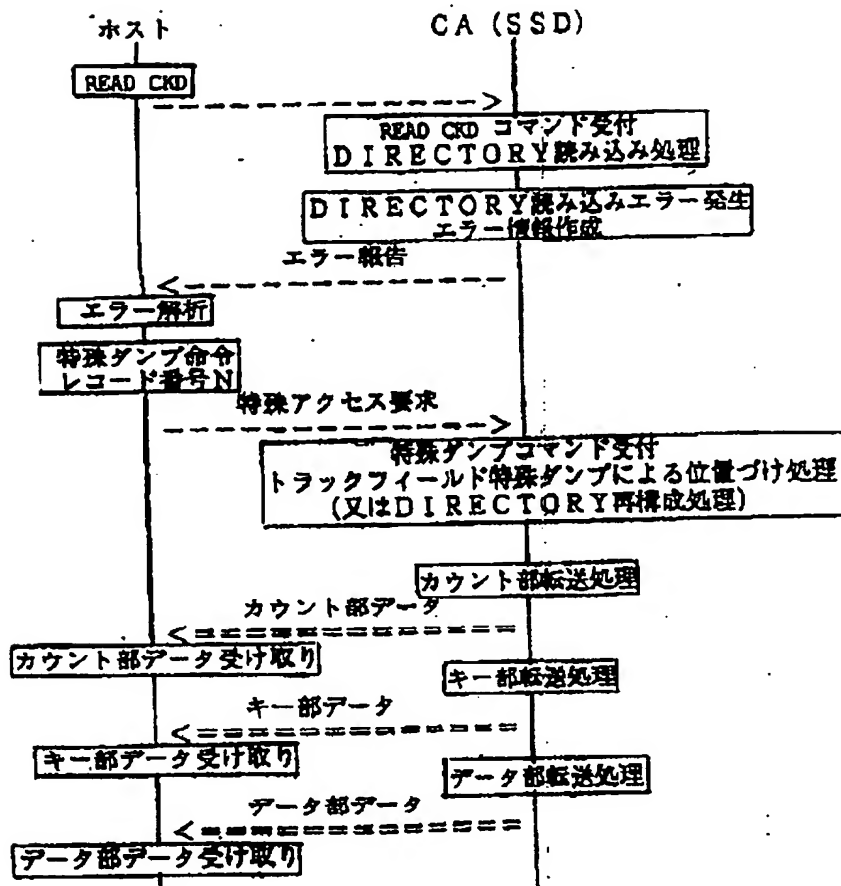


(33)

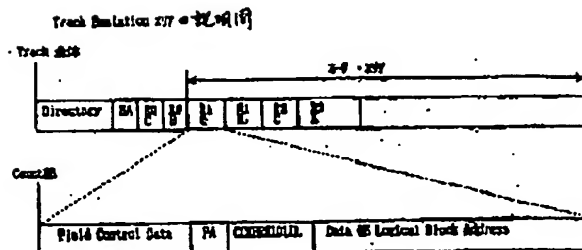
特開平7-311661

【図47】

トラック特殊アクセス法を用いた
ホストアクセスの説明図



【図52】

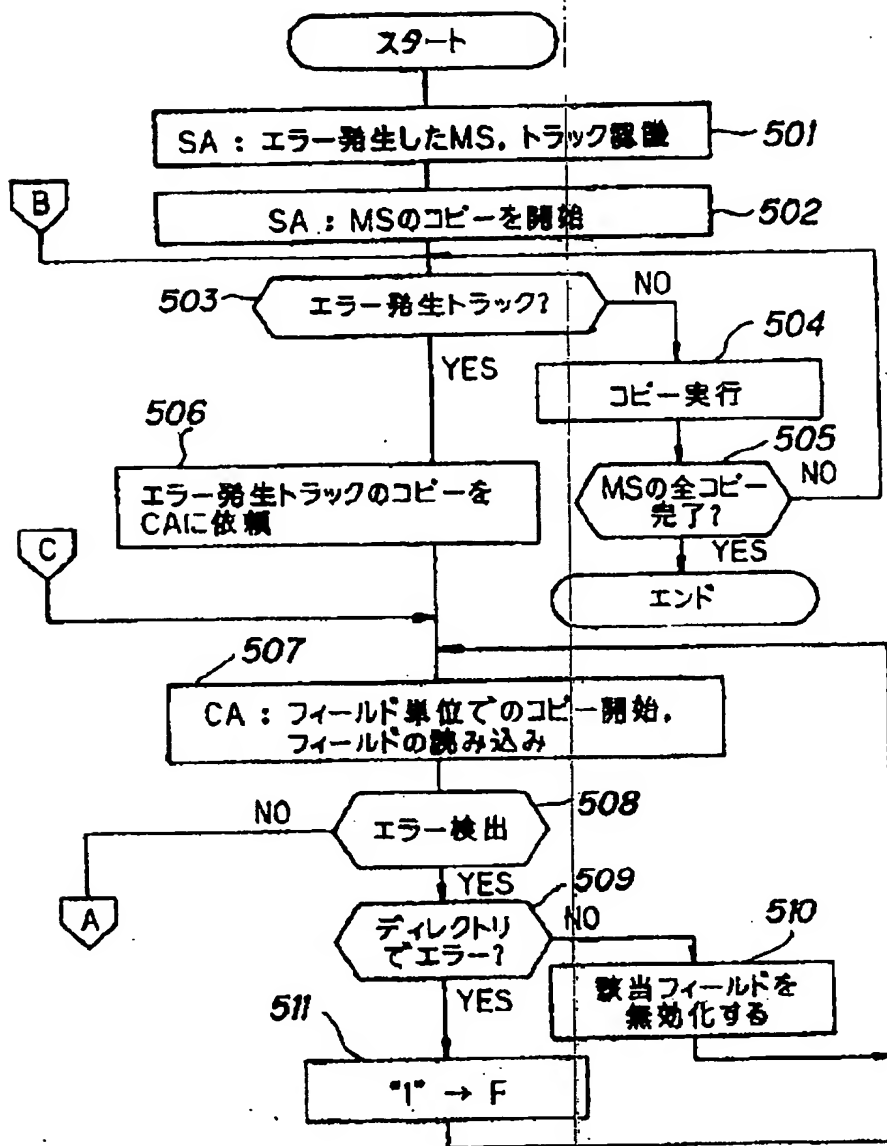


(34)

特開平7-311661

【図48】

半導体メモリにエラーが発生した時のコピー処理フロー(その1)

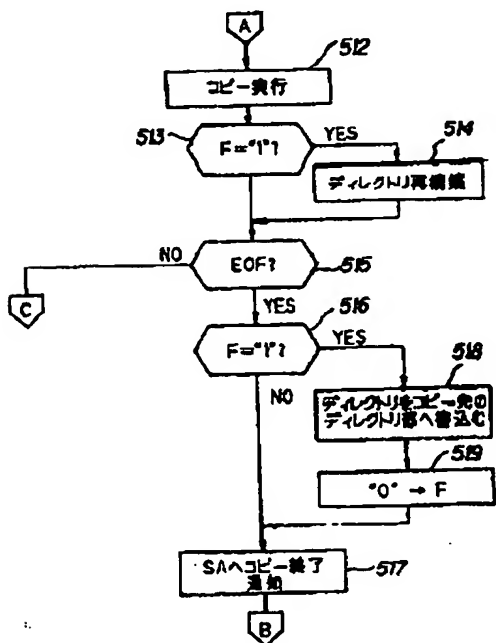


(35)

特開平7-311661

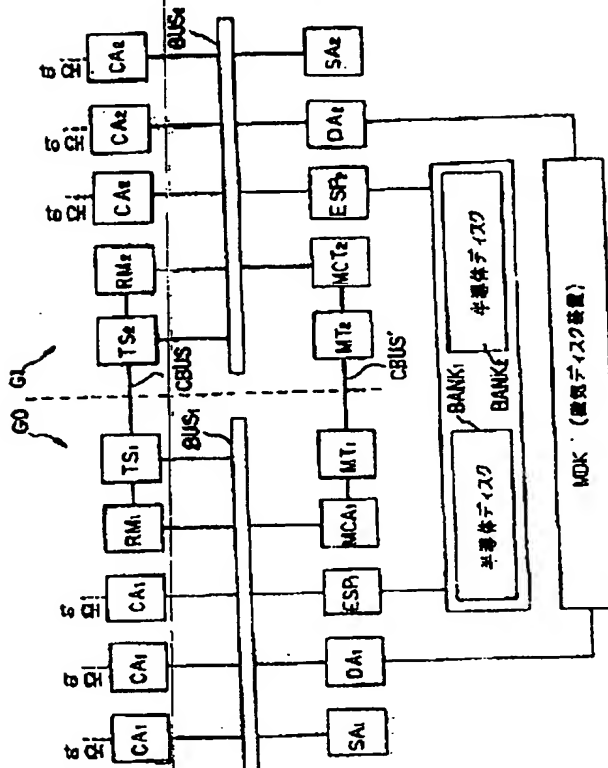
【図49】

半導体メモリエラーが発生した時のエラー処理フロー(その2)



【図50】

半導体ディスク装置の構成図(第4実施例)



【図53】

24-4-4 の Logical Block (LBK) 格納例

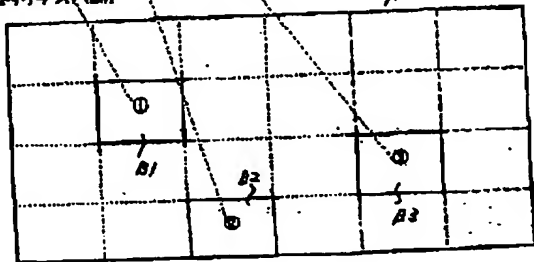
24-4-4 圧縮前



24-4-4 圧縮後

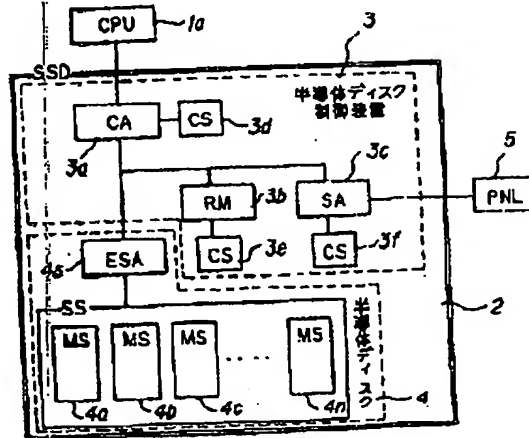


24-4-4 格子格納



【図57】

従来の半導体ディスク装置の構成



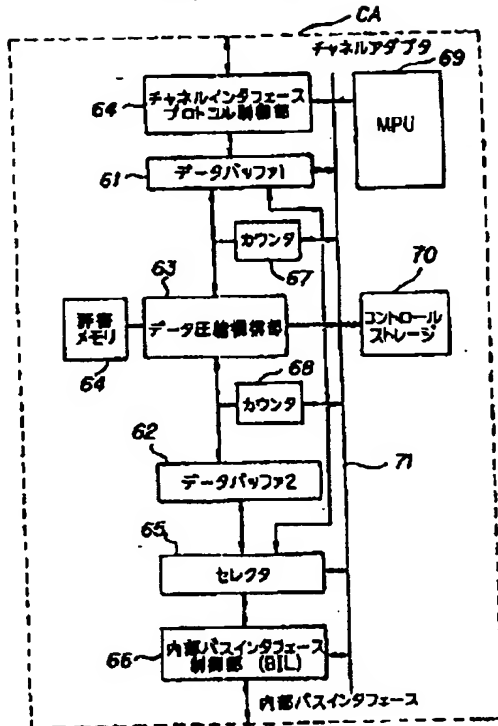
6.

(36)

特開平7-311661

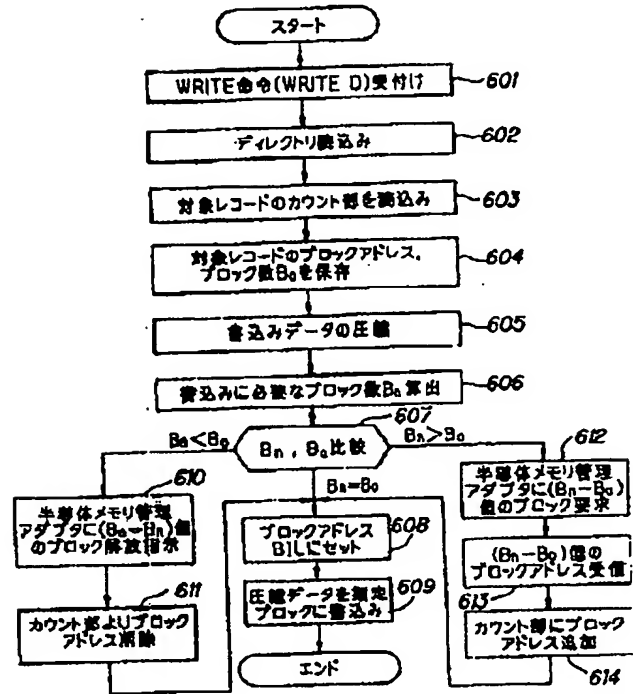
【図54】

チャンネルアダプタの構成図



【図56】

レコード書きの処理フロー



フロントページの続き

(72)発明者 西 英俊
神奈川県横浜市港北区新横浜2丁目4番19
号 株式会社富士通プログラム技研内